PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07K 14/47, 16/18, A61K 38/17, G01N 33/68

A1

(11) International Publication Number: WO 97/01578

(43) International Publication Date: 16 January 1997 (16.01.97)

(21) International Application Number: PCT/US96/10958

(22) International Filing Date: 27 June 1996 (27.06.96)

(30) Priority Data:

60/000,674	29 June 1995 (29.06.95)	US
08/663,617	14 June 1996 (14.06.96)	US

(71) Applicant (*for all designated States except US*): MEDICAL BIOLOGY INSTITUTE [US/US]; 11077 North Torrey Pines Road, La Jolla, CA 92037 (US).

(72) Inventor; and
(75) Inventor/Applicant (*for US only*): OGATA, Ronald, T. [US/US]; 549 Caretta Way, Cardiff, CA 92007 (US).

(74) Agent: ALTMAN, Daniel, E.; 16th floor, 620 Newport Center Drive, Newport Beach, CA 92660 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD FOR IDENTIFYING PEPTIDES THAT AFFECT PROTEIN-PROTEIN INTERACTIONS AND COMPLEMENT-MODULATING PEPTIDES

(57) Abstract

A method of identifying peptides, peptide analogs and peptidomimetics that have a high probability of inhibiting, enhancing or mimicking the activity of a target protein by comparing amino acid sequences of related proteins to identify indel-associated sequences is disclosed. The amino acid sequences of peptides that modulate the activity of the complement system are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	GB	United Kingdom	MW	Malawi
AT	Austria	GE	Georgia	MX	Mexico
AU	Australia	GN	Guinea	NE	Niger
BB	Barbados	GR	Greece	NL	Netherlands
BE	Belgium	HU	Hungary	NO	Norway
BF	Burkina Faso	IE	Ireland	NZ	New Zealand
BG	Bulgaria	IT	Italy	PL	Poland
BJ	Benin	JP	Japan	PT	Portugal
BR	Brazil	KE	Kenya	RO	Romania
BY	Belarus	KG	Kyrgystan	RU	Russian Federation
CA	Canada	KP	Democratic People's Republic of Korea	SD	Sudan
CF	Central African Republic			SE	Sweden
CG	Congo	KR	Republic of Korea	SG	Singapore
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LR	Liberia	SZ	Swaziland
CS	Czechoslovakia	LT	Lithuania	TD	Chad
CZ	Czech Republic	LU	Luxembourg	TG	Togo
DE	Germany	LV	Latvia	TJ	Tajikistan
DK	Denmark	MC	Monaco	TT	Trinidad and Tobago
EE	Estonia	MD	Republic of Moldova	UA	Ukraine
ES	Spain	MG	Madagascar	UG	Uganda
FI	Finland	ML	Mali	US	United States of America
FR	France	MN	Mongolia	UZ	Uzbekistan
GA	Gabon	MR	Mauritania	VN	Viet Nam

METHOD FOR IDENTIFYING PEPTIDES THAT AFFECT PROTEIN-PROTEIN INTERACTIONS AND COMPLEMENT-MODULATING PEPTIDES

RELATED APPLICATION

This application claims priority to provisional application No. 60/000674, filed June 29, 1995 under 35 U.S.C. § 119(e).

GOVERNMENT RIGHTS

This invention was made with United States government support under grants R01 GM29831 and R01 GM36960 awarded by the National Institutes of Health. The government has certain rights in the invention.

FIELD OF THE INVENTION

The present invention relates to modulating protein activity, and specifically relates to a general method of identifying regions of contact in protein-protein binding complexes, peptides or peptide analogs that modulate these protein-protein interactions, and peptides or peptide analogs that modulate complement activity.

BACKGROUND OF THE INVENTION

A fundamental component of essentially all biological processes is the specific interaction of proteins with other molecules, in particular other proteins. Specific protein-protein interactions are essential for cellular maintenance, regulation, reproduction and death. Protein-protein interactions are required for basic processes such as intercellular adhesion and communication, signal transduction, and gene replication, expression, and regulation. Specific protein-protein interactions are also critical in specialized processes such as blood clotting and the immune system, including the complement system. Thus, the ability to modulate specific protein-protein interactions is a key to modulating essentially all biological processes.

Protein-protein binding interactions occur at specific regions of contact on the protein surfaces and identification of the amino acid residues in these contact regions is important for controlling the binding interaction. For example, knowledge of the residues known to be involved in a contact site may be used to synthesize a relatively short peptide corresponding to the protein segment that includes these residues. This "interface" peptide may be used directly to compete with and thus inhibit that specific protein-protein binding interaction. Interactions that have been successfully inhibited in this manner include those between the herpes virus protein Vmw65 and the Oct-1 protein (Haigh, A. et al., 1990, *Nature* 344:257-259), the subunits of ribonucleotide reductases from viral and mammalian sources (Dutia, B.M. et al., 1986, *Nature* 321:439-441; Cohen, E.A. et al., 1986, Nature 321:441-443; Cosentino, G. et al., 1991, Biochem. Cell. Biol. 69:79-83), the subunits of the protease and reverse transcriptase of human immunodeficiency virus (HIV) (Zhang, Z.-Y. et al., 1991, *J. Biol. Chem.* 266:15591-15594; Babe, L.M. et al., 1992, *Prot. Sci.* 1:1244-1253; Schramm, H.J. et al., 1993, *Biochem. Biophys. Res. Commun.* 194:595-600; Divita, G. et al., 1994, *J. Biol. Chem.* 269:13080-13083), and the catalytic and accessory subunit, UL42, of herpes virus DNA polymerase (Digard, P. et al., 1995, *Proc. Natl. Acad. Sci. USA* 92:1456-1460).

An interface peptide may also be used indirectly to elicit an antibody which binds specifically to the intact protein. Such an antibody may be used to inhibit the protein-protein binding reaction. An interface peptide may be used as a structural model for design of a modified peptide, a peptide analog (Bianchi, E. et al., 1995, *J. Mol. Biol.*

247:154-60), or peptidomimetic inhibitor (McDonnell, J.M. et al., 1996, *Nature Structural Biol.* 3(5):419-426; Nakanishi, H. et al., 1993, *Gene* 137:51-56), as has been accomplished for a ribonucleotide reductase protein (Liuzzi, M. et al., 1994, *Nature* 372:695-698). If the protein-protein interaction is between a receptor and its ligand, an interface peptide may inhibit or substitute for the normal receptor-ligand interaction, and hence attenuate or mimic the normal cellular response depending on the signalling mechanism of the ligand-receptor binding.

Currently used methods for identifying residues at the protein-protein interface typically include direct analysis of the three-dimensional structure of the protein-protein complex (e.g., X-ray crystallography or NMR spectroscopy), or systematically testing the binding activity of protein fragments or natural or engineered variants of these proteins. In the studies listed above, where synthetic peptides were used to block protein-protein interactions, crystallographic data and genetically engineered modified proteins were used to identify interface residues in all cases except the herpes virus ribonucleotide reductase; in that case, a peptide designed to elicit an antibody response was fortuitously found to inhibit subunit association as well (Dutia, B.M. et al., 1986, *Nature* 321:439-441; Cohen, E.A. et al., 1986, *Nature* 321:441-443). All of the currently used methods are laborious and relatively unpredictable with regard to identifying interface peptides successfully, especially for proteins larger than a few hundred amino acid residues in length. X-ray crystallography requires crystallization of the protein complex in a form that can be analyzed. NMR spectrometry is limited by the size of the complex that can be analyzed. Methods involving modified proteins and their fragments require extensive testing, especially for larger proteins and without guidelines for identifying likely regions of contact.

The complement system, part of the mammalian humoral immune system, lyses microorganisms and infected cells by forming holes in the their plasma membranes. The complement system consists of more than twenty plasma and membrane-bound proteins which interact to trigger and modulate complement activity. While normal complement function is essential for health, regulation of the effects of complement activation is also important. For example, uncontrolled complement activation can lead to adverse symptoms such as continuous inflammatory reactions in a variety of diseases (Vogt, W., 1985, *Trends Pharm. Sci.* 6: 114-119). Also, complement is the primary mediator of the hyperacute rejection of xenogeneic transplants which limits the clinical and research usefulness of xenotransplantation (Ryan, U.S., 1994, *Xeno* 2: 19-22; Platt, J.L. et al., 1990, *Immunol. Today* 11: 450-456). Indeed, xenotransplantation as a viable surgical procedure will certainly require drugs that limit complement activation. Furthermore, therapeutic treatments of the reperfusion injury associated with myocardial infarction will likely include modulation of the effects of complement activation (Homeister, J.W. et al., 1994, *Annu. Rev. Pharmacol. Toxicol.* 34:17-40).

Despite a clear need for complement modulators, most of the currently available methods for inhibiting or depleting complement are not suitable for clinical applications because of their toxicity, undesirable side-effects or lack of efficacy in plasma (Vogt, W., 1985, *Trends Pharm. Sci.* 6:114-119). One complement inhibitor being clinically tested is soluble complement receptor type 1 (sCR1), which can suppress hyperacute rejection in xenotransplantation (Weisman, H.F. et al., 1990, *Science* 249:146-151; Pruitt, S.K. et al., 1994, *Transplantation* 57:363-370). sCR1 inhibits the complex C3 and C5 activating enzymes (convertases) by facilitating their dissociation. Although effective,

sCR1 is a large (240 kDa) protein and there is a need for small peptide inhibitors which are more easily synthesized, chemically modified, or mimicked by small organic peptidomimetic molecules (Nakanishi, H. et al., 1993, *Gene* 137:51-56).

The complement proteins C3, C4, and C5 are excellent targets for functional intervention by complement inhibitors because they play central roles in activation and regulation of this system. Although the three proteins have distinct functions, their amino acid sequences are closely related, and they are encoded by genes that doubtless evolved from a common ancestor (Campbell, R.D. et al., 1988, *Ann. Rev. Immunol.* 6:161-195). Because of their close structural relationship, the C3, C4 and C5 proteins form a protein family. All three proteins interact with a number of other proteins. In particular, C3 binds to or transiently interacts with more than a half-dozen soluble proteins including another molecule of itself during complement activation and attenuation, and to a similar number of cell-surface-bound proteins, which mediate immune-clearance, inflammatory, and complement regulatory activities. Researchers have identified the locations of sites within C3, C4 and C5 that interact with other complement proteins, including those in C3 that are recognized by inactivating proteases and their cofactors, cell-bound receptors and catalytic subunits in the complex complement convertases (Alsenz, J. et al., 1992, *Dev. Comp. Immunol.* 16:63-76).

Insertion/Deletion Sequences in Protein Families

During the evolution and divergence of proteins, individual members of a protein family undergo insertion or deletion of amino acid residues by corresponding insertion or deletion of DNA sequences in the genes encoding these proteins. This process results in length polymorphisms in the protein family. Because a deletion in one member of a family is equivalent to an insertion in others, these insertion/deletions have been referred to as "indels" (Kruskal, J.B., 1983, *in* Time Warps, String Edits and Macromolecules: Theory and Practice of Sequence Comparison, D. Sankoff and J.B. Kruskal, ed., pp. 1-44).

The locations of indels are revealed when amino acid sequences are aligned to maximize their sequence identity. Alignments of relatively short sequences can be carried out manually. Computer programs such as PILEUP (Genetics Computer Group, Madison, WI) and CLUSTAL W (Thompson, J.D. et al., 1994, *Nuc. Acids Res.* 22: 4673) are useful for aligning multiple long sequences. In aligning related sequences of different lengths, gaps are introduced into one or more family members to optimize the alignment of the total sequences. These gaps are indels.

Indels have been shown to occur in portions of the amino acid sequence of a protein that appear on the surface of the protein in its native folded state. Generally indels are short (1 to 5 residues) and occur at the protein surface as reverse turns or coils within loops rather than within secondary structural elements (α-helices and β-strands), because these properties minimize perturbations of the core protein structure (Pascarella, S. & Argos, P., 1992, *J. Mol. Biol.* 224: 461-471; Sibanda, B.L. & Thornton, J.M., 1993, *J. Mol. Biol.* 229: 428-447). Mutagenesis studies have experimentally confirmed that proteins are relatively tolerant of insertions within surface loops (e.g., see Freimuth, P.I. et al., 1990, *J. Biol. Chem.* 265: 896-901). Even within secondary structural elements, insertions can be tolerated if they occur at the protein surface (Betton, J.-M. et al., 1993, *FEBS Lett.* 325: 34-38).

SUMMARY OF THE INVENTION

This patent application describes a method for identifying likely regions of protein-protein contact based on comparison of protein sequences and describes examples of the use of the method for designing peptides that inhibit or enhance the action of the mammalian complement system. The method relies on identification of indels generally comprising selecting amino acid sequences for at least two proteins that contain similar amino acid sequences in at least a portion of the sequences and wherein one sequence is that of a target protein. The similar sequences include identical and/or conserved amino acid residues, and the similar sequences are aligned by matching the identical and/or conserved amino acid residues. Using the aligned sequences, sites are identified that contain insertions and/or deletions of amino acid residues in one protein relative to another protein, wherein the insertions and/or deletions define an indel. An indel-associated peptide sequence can either span or flank an indel. Indel associated peptides include peptides of about 4 to about 20 amino acid residues in length that are located within 30 amino acid residues, or preferably within 20, 15, 10, 9, 8, 7, 6 or 5 or less amino acid residues of an indel identified in the amino acid sequence of a target protein.

According to one aspect of the invention, there is provided a method for identifying molecules that affect biological activity of a target protein. The method includes the steps of obtaining information regarding the location of an indel in an amino acid sequence of a target protein, obtaining a peptide fragment of the target protein, the peptide fragment having a sequence that is located in the amino acid sequence of the target protein within 30 amino acids or less of the indel, or obtaining a peptidomimetic or peptide analog of the peptide fragment, and screening the peptide fragment, the peptidomimetic, or the peptide analog for its affect on biological or biochemical activity of the target protein. In a preferred embodiment, the screening step includes analyzing for modulation of protein activity, inhibition of protein activity, activation or potentiation of protein activity, competition for binding to a protein, binding to a protein or ligand, substitution for a substrate of the target protein, substitution for a ligand of the target protein, or making an anti-peptide antibody capable of modulating a biological activity of the target protein. In one embodiment of the method, the peptide fragment, the peptidomimetic, or the peptide analog directly affects the target protein whereas in another embodiment the peptide fragment, the peptidomimetic, or the peptide analog indirectly affects the target protein. Another preferred embodiment further comprises the step of synthetically constructing a peptide, peptide analog, or peptidomimetic that affects the biological or biochemical activity of the target protein. One preferred embodiment is a method for making a pharmaceutical composition, including the step of obtaining a molecule identified as having biological or biochemical activity in accordance with the steps of this method and combining the molecule with a pharmaceutically-acceptable carrier. This method may further include the step of packaging the molecule in unit-dosage form. In a preferred embodiment of the method, the target protein is a protein of the mammalian complement system. Preferably, the target protein which is a protein of the mammalian complement system is C2, C3, C4, C5 or Factor B. In a prefered embodiment, the peptide fragment has a sequence of about 4 to about 20 amino acid residues in length. In another embodiment, the peptide fragment has a sequence of about 10 to about 20 amino acid residues in length. In one embodiment, the peptide fragment has a sequence of about 4 to about 15 amino acid residues in length. In another embodiment, the peptide fragment has

a sequence of about 5 to about 18 amino acid residues in length. In one embodiment, the peptide fragment is located within about 20 amino acid residues of an indel. In another embodiment, the peptide fragment is located within about 15 amino acid residues of an indel. In yet another embodiment, the peptide fragment is located within about 12 amino acid residues of an indel. In one embodiment, the peptide fragment is located within about 10 amino acid residues of an indel. In another embodiment, the peptide fragment is located within about 9 amino acid residues of an indel. In yet another embodiment, the peptide fragment is located within about 8 amino acid residues of an indel. In one embodiment, the peptide fragment is located within about 7 amino acid residues of an indel. In one embodiment, the peptide fragment is located within about 6 amino acid residues of an indel. In other embodiments, the peptide fragment is located within about 5 or less amino acid residues of an indel or has a sequence that spans the indel or has a sequence located within the indel.

According to another aspect of the invention, there is provided a method of identifying interface peptides for a target protein, including the steps of identifying the location of an indel in the amino acid sequence of a target protein, selecting an amino acid sequence from the target protein sequence overlapping or located within about 30 amino acid residues or less of an amino- or carboxyl-terminus of an indel, obtaining a molecule that is a peptide having the selected amino acid sequence, a peptide analog of the selected amino acid sequence, or a peptidomimetic of the selected amino acid sequence, and evaluating the peptide, peptide analog or peptidomimetic in an assay to measure a change in activity of the target protein, wherein the change in activity is mediated directly or indirectly by the peptide, peptide analog or peptidomimetic. In one embodiment of the method, the peptide having the selected amino acid sequence is about 4 to about 20 amino acid residues in length. In an embodiment of the method, the peptide having the selected amino acid sequence is about 10 to about 20 amino acid residues in length. In another embodiment, the peptide having the selected amino acid sequence is about 4 to about 15 amino acid residues in length. In another embodiment, the peptide having the selected amino acid sequence is about 5 to about 18 amino acid residues in length. In one embodiment, the selected amino acid sequence has an amino- or carboxyl-terminal residue within about 20 amino acid residues of one terminus of the indel. In another embodiment, the peptide is located within about 15 amino acid residues of an amino- or carboxyl-terminus of an indel. In one embodiment, the peptide is located within about 10 amino acid residues of an amino- or carboxyl-terminus of an indel. In another embodiment, the peptide is located within about 9 amino acid residues of an amino- or carboxyl-terminus of an indel. In yet another embodiment, the peptide is located within about 8 amino acid residues of an amino- or carboxyl-terminus of an indel. In one more embodiment, the peptide is located within about 7 amino acid residues of an amino- or carboxyl-terminus of an indel. In another embodiment, the peptide is located within about 6 amino acid residues of an amino- or carboxyl-terminus of an indel. In one embodiment, the peptide is located within about 5 amino acid residues or less of an amino- or carboxyl-terminus of an indel. In a preferred embodiment, the method further includes the step of making antibodies to the peptide, peptide analog or peptidomimetic, wherein the antibodies are capable of modulating an activity of the targ protein. In a preferred embodiment, the change in activity of the target protein is a decrease in activity, an increase in activity, utilization of a substrate different than the substrate normally utilized by the target protein, or binding to a ligand differently than the ligand binding activity

ordinarily demonstrated by the target protein. Preferably, the target protein is a protein of the mammalian complement system and most preferably is C2, C3, C4, C5 or Factor B.

According to another aspect of the invention, there is provided a peptide for modulating activity of the complement system of a mammal, comprising a sequence of about 4 to about 25 amino acid residues that occurs in an amino acid sequence of a mammalian complement protein, the peptide having an amino acid sequence in which an amino- or carboxyl-terminal residue is located within about 15 amino acid residues of an indel of the mammalian complement protein. In one embodiment, the peptide modulates activity of the mammalian complement system by directly or indirectly inhibiting an activity of a protein of the mammalian complement system. In another embodiment, the peptide modulates activity of the mammalian complement system by directly or indirectly enhancing an activity of a protein of the mammalian complement system. In a preferred embodiment, the indel occurs within the amino acid sequence of a C2, C3, C4, C5 or Factor B protein. In another embodiment, the peptide further includes another molecule attached to the peptide. Another embodiment is an antibody that specifically recognizes such a peptide with another molecule attached to the peptide. Another preferred embodiment is a peptide analog or peptidomimetic molecule of a peptide according to this aspect of the invention. Another embodiment is an antibody that specifically recognizes such a peptide analog or peptidomimetic molecule. Another preferred embodiment is a pharmaceutical composition including a peptide, peptide analog or peptidomimetic molecule according to this aspect of the invention. A preferred embodiment is a pharmaceutical composition including an antibody that specifically recognizes a peptide, peptide analog or peptidomimetic molecule according to this aspect of the invention. In a prefered embodiment, the peptide is located within about 12 amino acid residues of an indel of the mammalian complement protein. In another embodiment, the peptide is located within about 10 amino acid residues of an indel of the mammalian complement protein.

According to another aspect of the invention, there is provided a peptide having a sequence of any one of: SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74 or SEQ ID NO:75.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 shows a prototypic indel, revealed by alignment of portions of the amino acid sequences of the human ("hum") and mouse ("mus") C3, C4, and C5 proteins, using the alignment program CLUSTAL W. The aligned amino acid sequences correspond to sequences around indel number 7 in FIG. 2A-2E and include humC3 (SEQ ID NO:1), musC3 (SEQ ID NO:2), humC4 (SEQ ID NO:3), musC4 (SEQ ID NO:4), humC5 (SEQ ID NO:5) and musC5 (SEQ ID NO:6).

FIG. 2A-2E shows the alignment of the complete amino acid sequences of the complement proteins human C3 (hC3; SEQ ID NO:7), mouse C3 (mC3; SEQ ID NO:8), human C4 (hC4; SEQ ID NO:9), mouse C4 (mC4; SEQ ID NO:10), human C5 (hC5; SEQ ID NO:11) and mouse C5 (mC5; SEQ ID NO:12) with indels of two or more residues labeled under the sequences; sequences selected for peptide synthesis are underlined and labeled as in Table 3. Double underlining indicates overlap of the peptide sequences.

FIG. 3A-3D shows the alignment of amino acid sequences of the DNA polymerases from herpes simplex virus (HSV) type 1 (HSV-1; SEQ ID NO:13), HSV type 2 (HSV-2; SEQ ID NO:14), HSV type 6 (HSV-6; SEQ ID NO:15), baculovirus (Baculo; SEQ ID NO:16), ictalurid herpes virus 1 (CCV; SEQ ID NO:17), Choristoneura biennis entomopoxvirus (ENTPOX; SEQ ID NO:18) and hepatitis B virus (HepB; SEQ ID NO:19). Asterisks (*) mark identical residues and periods (.) mark conserved residues.

FIG. 4A-4C shows the alignment of amino acid sequences of the human C3 (HuC3; SEQ ID NO:7), mouse C3 (MuC3; SEQ ID NO:8), rat C3 (RATC3; SEQ ID NO:20) and guinea pig C3 (GUIPIG; SEQ ID NO:21) proteins. Asterisks (*) mark identical residues and periods (.) mark conserved residues.

FIG. 5A-5C shows the alignment of amino acid sequences of retroviral polyproteins of the immunodeficiency viruses of humans (HIV-1, SEQ ID NO:22; and HIV-2, SEQ ID NO:23), Simian (SimianIV; SEQ ID NO:24), chimp (ChimpIV; SEQ ID NO:25) and cat (FelineIV; SEQ ID NO:26), of Rous sarcoma virus (RSV; SEQ ID NO:27), Moloney murine leukemia virus (MOLONEY; SEQ ID NO:28) and Friend murine leukemia virus (F-MULV; SEQ ID NO:29). Asterisks (*) mark identical residues and periods (.) mark conserved residues.

FIG. 6A-6B shows the alignment of part of the amino acid sequences of retroviral polyproteins of the immunodeficiency viruses of humans (HIV-1, SEQ ID NO:22; and HIV-2, SEQ ID NO:23), Moloney murine leukemia virus (MOLONEY; SEQ ID NO:28) and Friend murine leukemia virus (F-MULV; SEQ ID NO:29). Asterisks (*) mark identical residues and periods (.) mark conserved residues.

FIG. 7A-7B shows the alignment of amino acid sequences of ribonucleotide reductases from HSV-1 (SEQ ID NO:30), HSV-2 (SEQ ID NO:31), Epstein-Barr virus (EBV; SEQ ID NO:32), human (SEQ ID NO:33), vaccinia virus (Vaccinia; SEQ ID NO:34), mouse (Mus; SEQ ID NO:35), yeast (Yeast; SEQ ID NO:36), *Escherichia coli* (Coli; SEQ ID NO:37) and *Hemophilus influenzae* (H. infl.; SEQ ID NO:38). Asterisks (*) mark identical residues and periods (.) mark conserved residues.

FIG. 8A-8B shows the alignment of amino acid sequences of herpes virus Vmw65 (VP16) proteins for HSV-1 (HSV TYPE1/F; SEQ ID NO:39), HSV-2 (HSV TYPE2/HG52; SEQ ID NO:40), bovine virus (BvHV Type1/P8-2; SEQ ID NO:41), Varicella-Zoster virus (Var-ZosV/Dumas; SEQ ID NO:42) and equine viruses (EqHV Type4, SEQ ID NO:43; and EqHV Type1/AB4P, SEQ ID NO:44). Asterisks (*) mark identical residues and periods (.) mark conserved residues.

FIG. 9 graphically shows inhibition of complement activity (CH50) by the C3 peptides II-4 (●), II-5 (C, □) II-8C (⊖, ⊏, ◇) and III-11 (+, ×) with K_i values for the dotted line (....) being 4 X 10^{-5}, for the uniformly dashed line (----) being 6 X 10^{-5} and for the solid line (____) being 2 X 10^{5}.

FIG. 10 shows the peptide concentrations giving 50% inhibition of complement function as measured by hemolysis for the individual peptides listed under the bars on the X-axis.

FIG. 11A-11B shows the alignment of amino acid sequences of human C2 (humC2; SEQ ID NO:45), mouse C2 (musC2; SEQ ID NO:46), human factor B (humBf; SEQ ID NO:47), mouse factor B (musBf; SEQ ID NO:48) and zebrafish factor B (ZebBf; SEQ ID NO:49).

FIG. 12 graphically shows the inhibition of complement hemolytic activity by the C2 peptide E1 for concentrations of peptide indicated on the X-axis used without serum (×) or with serum (▵, ◇ and □).

DETAILED DESCRIPTION OF THE INVENTION

A general method is disclosed for identifying portion of proteins that are involved in specific protein-protein interactions or contacts with other molecules. Identifying these interactive portions of proteins constitutes a critical step in designing small molecule inhibitors of protein-protein interactions. This method involves the analysis of only the primary structures of related proteins and has been demonstrated to be effective in predicting the amino acid sequences of inhibitory peptides for the complement system.

In the absence of a three-dimensional structure for a protein, identifying the amino acid residues that constitute a protein's interactive sites presents a daunting challenge, particularly for large proteins. For example, for the complement proteins C3, C4 and C5, this requires identifying such sites among greater than 1600 amino acid residues in each protein.

Nevertheless, using other more time-consuming methods, researchers have identified sites within C3, C4 and C5 that interact with other complement proteins, including those in C3 that are recognized by inactivating proteases and their cofactors, cell-bound receptors and catalytic subunits in the complex complement convertases (Alsenz, J. et al., 1992, *Dev. Comp. immunol.* 16:63-76). A key step in the development of the present method was the inventor's observation that almost all of these reported interactive sites span or lie within a few amino acid residues of primary structural features known as indels. The method described here for identifying regions of specific protein-protein contact relies on analysis of length polymorphisms and on determination of the locations of indels in protein families. This general method provides a means for identifying interactive sites on proteins with a high probability of success.

For purposes of this application, certain terms are defined as follows.

"Identity" means that all of the amino acid residues at a single position are identical when two or more protein sequences are aligned. Two or more protein sequences may have only limited regions of identity and gaps may be introduced into one or more sequences to aid in alignment of identical residues.

"Conserved" amino acids means that non-identical amino acid residues share sufficient chemical similarity that they can be considered to form a functional group. That is, when amino acid residues for two or more protein sequences are compared, a single position of alignment may contain non-identical amino acids but the residues may be conserved for that position and functionally similar because of their chemical and/or structural similarity. For purposes of this application, the following groups of amino acids (indicated by their three letter code followed by their one letter code in parentheses) form groups of conserved amino acids:

Phe (F) and Tyr (Y);

Ile (I), Leu (L), Val (V) and Met (M);

Ala (A), Ser (S) and Thr (T);

Asn (N), His (H), Arg (R), Lys (K) and Gln (Q); and

Asn (N), Asp (D), Glu (E) and Gln (Q).

For comparison of two or more protein sequences, conserved amino acids for any given position means that all of the amino acid residues at a single position are members of a single group of conserved amino acids although any

combination of amino acids within that conserved group is permissible. The definition of conserved residues used here follows closely that defined by the PAM250 matrix (Dayhoff, M.O., et al. 1978. *In* Atlas of Protein Sequence and Structure, Vol. 5, suppl. 3 (Dayhoff, M.O., ed) p. 345, NBRF, Washington, DC).

"Alignment" of amino acid sequences allows for insertion of spaces in one or more of the sequences of amino acid residues to maximize the number of positions having identical and/or conserved residues in the compared protein sequences using methods well known in the art. For purposes of alignment of more than two sequences, the identical or conserved residues do not have to be present in all of the sequences for any particular location (i.e., an identical or conserved residue may be identified even if it is identical or conserved for that locus only for two of the compared sequences). It will be understood that in an alignment of two or more similar sequences there can be a mixture of identical, conserved and non-identical nonconserved amino acid residues in the aligned sequences. The comparisons resulting in alignment can be done manually or using computer software programs such as, for example, the CLUSTAL W and PILEUP programs.

An "interface peptide" is a peptide that includes an amino acid sequence that occurs in a protein at or near a specific protein-protein contact region. An interface peptide can specifically inhibit a protein-protein interaction by acting as a competitive inhibitor of the native protein or can serve as a model for making a modified peptide, a peptidomimetic or anti-peptide antibodies.

An "indel" is a region of amino acid sequence that includes all "deleted" residues as shown in at least one amino acid sequence when aligned with another amino acid sequence(s), where the sequences share overall similar sequences of amino acid residues. That is, the compared sequences contain some identical or conserved amino acid residues occurring at the same or approximately the same positions in the compared protein sequences. An indel spans the farthest extent of deleted (or inserted) amino acid residues as discussed in detail below and shown in FIG. 1. In cases where insertion/deletions are less than ten residues apart (e.g., see indel 13 in FIG. 2A-2E), the entire segment containing the multiple insertions/deletions is regarded as a single indel. Indels can also occur at the amino- and carboxyl-terminal ends of a protein.

A "divergent region" is a portion of an amino acid sequence which shows dissimilarities in the amino acid residues of at least two members of a protein family whose amino acid sequences are compared. The divergent region is bordered by segments of amino acid sequences having high sequence conservation among the family members. An indel is the focal point of a divergent region. That is, a divergent region is a segment of the amino acid sequence including the indel that is relatively polymorphic among the compared sequences and is bounded by highly conserved regions.

When used to refer to the location of an amino acid residue of a peptide relative to an indel, "located within" means that at least one residue of the peptide is contained within the length of the sequence from at least one end of the indel. For example, if a peptide is located within 30 amino acids of an indel, at least one residue of the peptide is located within the sequence defined by residues 1 to 30 away from the indel. That is, at least one residue of the peptide is located between the end of the indel and the residue located 31 residues from the end of the indel.

Referring to FIG. 1, an indel is shown as the region that includes all "deleted" residues, shown by spaces occupied by asterisks (*), with the indel being the focal point of a divergent region among the protein family members which is bordered by segments having high sequence conservation among the family members. The spaces (shown by asterisks (*)) are introduced into the amino acid sequences to maximize the alignment of the flanking amino acid residues of the related proteins. As defined herein, the indel shown spans the farthest extent of deleted (or inserted) amino acid residues, shown by the arrow symbols ("<" and ">"), with "N" and "C" under the arrow symbols identifying the amino and carboxyl termini of the indel, respectively. In FIG. 1, the position marked with an arrow (↓) is defined as being six residues away from the amino terminal end of the indel. This figure compares small portions of the human and mouse amino acid sequences of the complement proteins C3, C4 and C5. Because of the multiple interactive properties of these proteins and their close sequence similarity, this protein family is used to demonstrate the utility of the disclosed method for identifying regions of proteins that are involved in interactions with other proteins.

FIG. 2A-2E shows the alignment of the complete amino acid sequences of human and mouse C3, C4, and C5 proteins, which reveals the positions of indels in this protein family. The residues shown in lower case letters in C3 and C4 have previously been identified by other methods as interactive sites in these proteins. Some of these sites are listed in Table 1 with summaries of the functional features, thus showing that the reported interactive sites are generally indel-proximal as defined herein.

Table 1. Interactive sites near indels in the C3/C4/C5 protein family.

Indel No.	Functional Feature
13-14*	Activating protease cleavage site in C3, C4, and C5; located between indels 13 and 14 are the binding sites for C3 or its proteolytic fragments with the following proteins: Factor B Factor H Complement receptor 1 (CR1) Complement receptor 2 (CR2) Complement receptor 3 (CR3) Membrane cofactor protein (MCP)
16*	Specific protease cleavage sites in C3b, giving C3dK/dg Factor B binding site
21*	Factor H binding site CR2 binding site Cross-linking site of C4b to C3b thioester in the C5 convertase
22	Specific inactivating proteolytic sites in C3b and C4b
26	Properdin binding site

* Indels that occur at intron/exon junctions.

The occurrence of indels primarily as loops on protein surfaces led me to hypothesize that indels are in general involved in the interactions of proteins with other molecules. Hence, in the absence of an NMR or X-ray crystal structure of the protein, indels provide focal points for identifying amino acid residues within a protein sequence that are likely to be involved in interactions with other molecules. The disclosed method for identifying likely sites of protein-protein interactions is based on alignments of the amino acid sequences of related proteins and identification of indels in the protein family. Protein sequences in databanks such as GENBANK and SWISS-PROT are readily available to those skilled in the art for use with this method which provides a rapid and effective approach to defining sites of protein-protein interactions compared to currently available procedures. This method has been used to identify linear amino acid sequences of peptides that significantly affect (inhibit or enhance) complement activity.

The method for identifying protein-protein and protein-substrate interactive sites includes the steps of aligning the amino acid sequence of the target protein of interest with the sequences of one or more closely related proteins, and identifying indels in the protein family using this alignment. Because indels usually mark protein segments at the surface of a naturally folded protein, they also mark regions of potential protein interactive sites. Although only a subset of indel-proximal regions may actually be involved in a particular protein-protein interaction, focusing on indels greatly reduces the task of finding the desired contact site(s) compared to currently used methods. Involvement of a particular indel-proximal region in protein-protein binding may be tested in several ways. For example, synthetic peptides containing indel-associated sequences may be constructed and tested for modulating a particular protein-protein interaction by measuring the activity of the protein of interest, in the presence and absence of peptide, in an assay in which that interaction is obligatory. Alternatively, the peptides may be used to elicit anti-

peptide antibodies which are similarly tested. The peptide may also be used as a structural model for construction of modified peptides or peptidomimetics which are tested in the same manner. Peptide-based tests for involvement of a particular indel-proximal region may fail to reveal interface peptide activity even if the region is involved in a protein-protein interaction because the assay may require the peptide to be in a three-dimensional conformation like that of the corresponding segment within the native protein. Additional methods for indel testing include engineering proteins with sequence modifications at the indel. No matter which assay is chosen, focusing on indels greatly simplifies the search for interactive sites.

Because a deletion in one member of a family is equivalent to an insertion in another and because the size of an indel may vary from one to about 50 amino acid residues, the location of an indel is defined as follows. From the point of view of an amino acid sequence containing a deletion relative to another protein (defined as the insertion-containing member) with which it shares a similar amino acid sequence, the indel site begins at the first residue that is missing in the deletion-containing member(s) relative to the insertion-containing member(s) and ends at the last residue that is missing in the deletion-containing member(s) relative to the insertion-containing member(s) (illustrated in FIG. 1). For a protein family in which an indel consists of a single amino acid insertion in one protein, the indel will be a single amino acid residue in length. For a protein family in which an indel consists of two or more residues, the indel will have amino and carboxyl termini (as indicated with the "N" and "C" in FIG. 1). The length of the indel is irrelevant to its identification or its predictive utility for identifying indel-associated peptides. Peptides associated with such indels may be proximal to either the amino terminus or the carboxyl terminus or may span the indel. Indel-associated peptides that flank an indel do not necessarily begin or end immediately adjacent to the amino- or carboxyl-terminus of the indel but can have an amino- or carboxyl-terminus within about 30 amino acid residues of the indel.

In cases where insertion/deletions are less than ten residues apart and contain identical or conserved amino acid residues between them (e.g., indel 13 in FIG. 2A-2E), the entire segment is regarded here as a single indel. Indels can also occur at the amino- and carboxyl-termini of a protein as indicated by length polymorphisms at the termini of related proteins (e.g., the carboxyl-termini of the protein sequences shown in FIG. 7A-7B).

The process of identifying indels involves selection of sequences to be aligned, for example, from databanks such as GENBANK. Selection of sequences may be based on functional similarities of proteins, genetic relatedness, random scanning of database sequences for related sequences or other well-known molecular biology methods or combinations of methods. The selected sequences are aligned usually by using computer programs (e.g., PILEUP or CLUSTAL W programs) but may also be aligned manually. Although this process adequately identifies the positions of most indels, it will be understood by those skilled in the art that the location of some indels may be somewhat unsystematic depending on the method (e.g., algorithm) used to align the amino acid sequences and on the particular protein family members selected. This variability is most pronounced in regions of limited sequence similarity. Hence, avoiding unnecessary sequence dissimilarity, by careful selection of the individual protein family member sequences to be aligned, helps to minimize ambiguities in indel locations. Even if the exact indel location is uncertain, the method nevertheless gives an approximate location of a potential protein interactive site.

In selecting individual protein family members for comparison, sequence similarity and the functional properties of the proteins should be considered. The proteins must have sufficiently similar sequences to provide regions of unequivocal sequence alignment but must contain the length polymorphisms caused by insertions and/or deletions. For example, in the alignment of DNA polymerases shown in FIG. 3A-3D, the sequence similarities among the proteins is poor. Thus, the number and precise locations of indels is ambiguous. In contrast, the alignment of the very similar C3 proteins from different species in FIG. 4A-4C shows only a single indel greater than one residue in length. Therefore, the alignment of C3 with C4 and C5 in FIG. 2A-2E is more informative of potential locations of interactive sites in C3 and functionally important sites that are unique for each member of this family.

Specific types of protein-protein interactions may be examined by the indel method by appropriate choices of sequences that are compared (e.g., species-specific protein-protein contact sites as shown in FIG. 4A-4C). Because it is known that some complement components are incompatible with components from other species (e.g., *see* Kai, S. et al., 1980, *J. Immunol.* 125:2409-2415), the single large indel in the C3 alignment of FIG. 4A-4C may identify an interactive site that mediates such a species-specific interaction. Similarly, alignments of related protein sequences from closely related viruses having different tropisms may identify tropism-related interactive protein sites. Alignment of sequences for homologous proteins from different tissues may reveal tissue-specific functional regions of the proteins.

The selection of proteins for alignment can alter the appearance of indels as shown in FIGS. 5 and 6, in which different subsets of retroviral polyproteins are aligned. As shown in FIG. 5A-5C, when all eight sequences are included in the analysis, the number and locations of indels are somewhat ambiguous. In contrast, when only four of these sequences are compared in FIG. 6A-6B the indels become more clearly defined. These alignments demonstrate that for the clearest results, the sequences chosen must be similar enough to provide regions of significant sequence similarity. One skilled in the art can readily select sequences for making meaningful indel comparisons with a minimum of experimentation.

The nature of the protein-protein interaction is another important consideration in using this method. For example, amino acid sequences involved in obligate interactions (e.g., the α and β subunits of hemoglobin) tend to show less sequence variation and length polymorphism than segments involved in transient protein-protein associations. Furthermore, inhibitor-mediated disruption of protein-protein contacts in obligate complexes is generally difficult because, unless the complex is kinetically relatively unstable, the inhibitor can act only if it is present before or during complex formation. Peptidomimetics that readily enter the cell may be used in these situations as is well known in the art.

One aspect of this invention is that indels are used as guides in designing synthetic interface peptides that modulate the interaction between two proteins. Although a practitioner skilled in the art may use personal discretion in selecting interface peptide(s), the following two guidelines are useful. First, utilize sequences contained within the region of sequence divergence surrounding each indel as described above and illustrated in FIG. 1. That is, use only indel and flanking divergent sequences bounded by highly conserved sequences. I picture the divergent region as being at or near the protein surface, and hence that its sequence can vary without greatly perturbing the protein

conformation and is available for interactions with other proteins. In contrast, the conserved residues are inferred to be invariant because they are closer to the protein core, taking part in intramolecular interactions that maintain the overall protein conformation and are physically less available for intermolecular interactions. Second, select peptides that correspond in sequence to segments within the divergent region that are intermediate in hydropathy (i.e., amphipathic peptides). That is, those peptides having hydropathy between about 0.25 and 0.65 when calculated as the mean of the amino acid parameters as described by Fauchere, J.L., et al., 1983, *Eur. J. Med. Chem.* 10:369. Hydropathy can be determined by any of a number of well known methods. As shown in Tables 3 and 4 (discussed below), most effective peptide inhibitors have intermediate hydropathies. Although not all peptides with intermediate hydropathies show inhibitory activity (e.g., peptide I-6), none of the strongly hydrophilic peptides (hydropathy less than 0.25) showed 50% or more inhibition in an *in vitro* assay at concentrations of less than or equal to 150 μM (discussed in detail in Example 3).

Using the indel method, one skilled in the art can select segments of the target protein sequence for designing and synthesizing peptides, peptide analogs and peptidomimetics. Indel associated peptides include peptides of about 4 to about 20 amino acid residues in length that are located within 30 amino acid residues, or preferably within 20, 15, 10, 9, 8, 7, 6 or 5 or less residues of an indel identified in the amino acid sequence of a target protein. In selecting the amino acid sequences for making peptides (and for designing peptide analogs or peptidomimetics of the peptides), certain combinations of peptide length and distance relative to the indel are preferred. A peptide length of about 20 amino acid residues, located within about 10 residues of an indel is preferred, and a peptide of about 18 amino acid residues, located within about 8 residues of an indel is particularly preferred. Most preferably, the selected peptide includes about 15 amino acid residues and is located within about 6 residues of an indel.

Once an interface peptide is identified using the method and guidelines, it may be modified with respect to length and sequence for maximum effect using techniques well known to those skilled in the art. Modification by inclusion of conserved residues, presumably from the core of the folded protein, may be desirable to stabilize the conformation of the peptide in a form similar to that of the native protein. Peptides may also be acetylated at their amino terminus and/or amidated at their carboxyl terminus. Standard solid-phase peptide synthetic techniques allow for essentially unlimited quantities of the synthesized peptide of interest to be chemically produced (e.g., see Erickson & Merrifield, The Proteins, 3rd Ed., Vol. 2, Chapt. 3, Academic Press, New York, 1976; and Merrifield & Barany, The Peptides: Analysis, Synthesis, Biology, Vol. 1, Chapt. 1, Gross & Meinenhofer, eds., Academic Press, New York, 1980).

Peptide analogs and peptidomimetics are molecules that are modeled on a known peptide and synthesized generally to force a desired conformation on a peptide or peptide-like molecule by introducing conformational restraints (McDonnell, J.M. et al., 1996, *Nature Structural Biol.* 3(5):419-426; Nakanishi, H. et al., 1993, *Gene* 137:51-56; Liuzzi, M. et al., 1994, *Nature* 372:695-698). That is, a successful peptide analog or peptidomimetic includes appropriate functional groups positioned on a relatively rigid molecular framework. Conformational restraints may be introduced in a variety of ways, all well known in the art, including using amino acid residues that display

strong conformational tendencies, covalently cyclizing a peptide backbone (e.g., by introducing a disulfide bond between two cysteine residues in a peptide sequence), and introducing bulky chemical groups as side chains or as terminal modifications to a peptide (e.g., introducing one or more $(C_6H_5)CH_2$ groups at the amino-terminus of a peptide).

Peptide analogs and peptidomimetics mav be designed in a variety of ways yielding molecules that mimic the peptide of interest. One type of analog consists of the amino acid sequence of the identified peptide of interest with cysteine residues added at the amino- and carboxyl termini of the peptide. The resulting peptide is then oxidized to join covalently the terminal cysteine sulphydryl side chains, producing intrachain disulfide bridge and forming a cyclic peptide. Moreover, the corresponding linear version of the peptide is easily produced by reducing the disulfide bond with a reducing agent such as dithiothreitol (DTT) or by reductive alkylation with vinylpyridine (McDonnell, J.M. et al., 1996, *Nature Structural Biol.* 3(5):419-426).

Peptide analogs may contain some or all D-amino acid residues substituted for the L-amino acid residues in the peptide sequence of interest. D-amino acid containing peptides are often preferred for *in vivo* use because of their greater serum stability compared to peptides containing L-amino acids. One type of D-amino acid containing peptide analog is a retro-enantiomeric peptide consisting of only D-amino acid residues occurring in the reverse order of the sequence of L-amino acids in the peptide of interest. Such a retro-D-amino acid peptide is an isomer of the corresponding L-amino acid peptide, having reversed peptide bond orientation but theoretically identical side-chain topology as in the corresponding L-amino acid peptide. (McDonnell, J.M. et al., 1996, *Nature Structural Biol.* 3(5):419-426; Guptsarma, P., 1992, *FEBS Lett.* 310:205-210; Goodman, M. & Chorev, M., 1979, *ACC Chem. Res.* 12:1-7).

Another type of peptide analog retains the amino acid composition of the peptide of interest but consists of a "scrambled" amino acid sequence having any one of the finite number of possible combinations for that amino acid composition. One type of peptidomimetic uses structures within the peptide sequence that mimic the reverse-turns that occur naturally in proteins. That is, the peptide includes a β-turn that causes the peptide chain to reverse direction, for example, by limiting the distance between the Cα of the first residue and Cα of the fourth residue to about 4 to 7 Å. Such peptides are readily produced by a macrocyclization reaction involving the use of an azetidinone as an activated ester during peptide synthesis (Nakanishi, H. et al., 1993, *Gene* 137:51-56). Similarly, peptides containing a δ-turn can also be synthesized to produce δ-turn peptidomimetics.

Combinations of any of these techniques mav also be used to produce peptide analogs and peptidomimetic molecules. For example, a peptide analog may be produced by using a scrambled sequence of the peptide of interest and introducing one or more D-amino acid residues into the peptide.

Synthesis of peptide analogs and peptidomimetics is well known in the art and generally involves: (1) molecular modeling of a known peptide sequence which can be performed using any cf a variety of molecular modeling computational chemistry methods (e.g., those available as Discover™ and Homology™ from Biosym Technologies Inc. or the PROCHECK program of Laskowski, R.A. et al., 1993, *J. Appl. Crystallogr.* 26:283-291), and (2) chemical synthesis using well known peptide synthesis or organic synthesis techniques. Synthesis of peptide

analogs or peptidomimetics generally uses modifications of solid-phase peptide synthesis protocols (see, e.g., the synthetic methods discussed in Nakanishi, H. et al., 1993, *Gene* 137:51-56 and McDonnell, J.M. et al., 1996, *Nature Structural Biol.* 3(5):419-426). For example, peptide analogs and peptidomimetics may be produced using standard protein synthetic chemistry on an automated system (e.g:, Applied Biosystems 430A peptide synthesizer) but including one or more D-amino acid residues or by modifying the synthesis to include a macrocyclization reaction as discussed above. The synthesized peptides are then readily purified by any of a variety of techniques such as reverse phase high pressure liquid chromatography (RP-HPLC). Screening of existing chemical libraries or natural products for molecules that mimic the peptide of interest is an alternative method of identifying peptidomimetics without synthesizing the molecules *de novo*.

Using the guidelines presented in this disclosure for identifying indel-associated peptides greatly aids one in producing peptides, peptide analogs or peptidomimetics with biological activity. Nonetheless, even similar peptides associated with an indel can vary considerably in their activity. As shown in Table 2, the inhibitory activities of four overlapping peptides that were selected as candidate inhibitors based on their proximity to indel 25 (see FIG. 2A-2E) varied from very effective (peptide III-8C) to relatively ineffective (peptides II8A and III8B) inhibitors, while one peptide (II-8) was difficult to chemically synthesize and purify. Despite these difficulties, the method permitted rapid identification of one interface peptide from four candidate sequences. Thus, the method permits relatively efficient identification of active interface peptides with a minimum of empirical testing compared to other methods of searching for interactive sites and functionally active peptides.

Table 2. Complement inhibitory activity of peptides having indel 25-proximal C3 segments.

SEQ ID NO:	PEPTIDE	SEQUENCE	IH_{50} (μm)
50	II-8	Ac-DATMSILDISMMTG	NT*
51	II-8A	Ac-DATMSILDISMMTG-NH_2	> 300
52	III-8B	Ac-DQDATMSILDISMM-NH_2	> 400
53	III-8C	Ac-SILDISMMTGFAPDT-NH_2	60

* Not tested ("NT") because the peptide synthesis failed.
Arrow (↑) shows location of indel.

The identification of candidate indel-proximal interface peptides is inexact but can readily be determined by a practitioner skilled in the art using the techniques described herein. Using this method, the amount of experimentation needed to confirm the *in vitro* or *in vivo* activity of potential interface peptides is significantly reduced because of the relatively high probability of predicting interface peptide sequences. In practice, the activity

of an individual peptide, peptide analog or peptidomimetic may depend on its three-dimensional conformation, its solubility in physiological conditions and its ability to be chemically or biologically synthesized and purified. These synthesis, purification and assay techniques, however, are well known in the art and the activity of individual peptides, peptide analogs and peptidomimetics can be readily determined using standard techniques.

The utility of the present method for identifying interface peptides is supported by previous studies that have identified interface protein sequences using other methods. That is, the indel method described here has been applied retrospectively to active interface peptides that were identified using other methods. Where interface peptides have been shown to inhibit a protein-protein interaction, the inhibitory peptides are indel-proximal as defined by the present method. These include interface peptides that have been used to inhibit (1) the HIV protease (Zhang, Z.-Y. et al., 1991, *J. Biol. Chem.* 266:15591-15594; Babe, L.M. et al., 1992, *Prot. Sci.* 1:1244-1253; Schramm, H.J. et al., 1993, *Biochem. Biophys. Res. Commun.* 194:595-600), (2) HIV reverse transcriptase (Divita, G. et al., 1994, *J. Biol. Chem.* 269:13080-13083), the (3) herpes virus and human ribonucleotide reductases (Dutia, B.M. et al., 1986, *Nature* 321:439-441; Cohen, E.A. et al., 1986, *Nature* 321:441-443; Cosentino, G. et al., 1991, *Biochem. Cell. Biol.* 69:79-83), and (4) the herpes virus Vmw65 transcriptional control protein (Haigh, A. et al., 1990, *Nature* 344:257-259). These inhibitory peptide sequences had been selected on the basis of X-ray crystallographic structural information, by fortuitous observation of inhibitory activity and analogy to related proteins, and by individual testing of modified proteins.

When the sequences of the above four proteins were aligned (using the CLUSTAL W program) with those of related proteins to display indels, in all cases the inhibitory interface peptides corresponded in sequence to segments near indels as illustrated in FIGS. 6 to 8. Only a single previously reported interface peptide inhibitor could not be unequivocally associated with an indel: the peptide of the herpes virus DNA polymerase identified by Digard, P., et al., 1995, *Proc. Natl. Acad. Sci. USA* 92:1456-1460. The sequence of this polymerase is shown aligned with sequences of related polymerases in FIG. 3A-3D which shows that the sequences chosen for comparison are too dissimilar to give reliable indel locations thus precluding association of the interface peptide with an indel. Hence, with one possible exception, in previously published cases where interface peptides have been used successfully, these peptides are indel-proximal as defined by the disclosed method. These observations, together with the success in identifying complement inhibitory peptides based on indel proximity described in the Examples below, support the validity of the indel method for identifying interface peptides.

The examples focus on identifying inhibitory peptides based on analysis of indels of the complement C3/C4/C5 and C2/factor B protein families. The C3/C4/C5 family has well-defined structural characteristics useful for confirming the efficiency of the method for active peptide identification. It will be understood by those skilled in the art that these nonlimiting Examples provide a foundation for identifying active peptides for other protein-protein interactions for which primary amino acid sequence information is known which would allow one to identify indels, such as, for example, the protein families illustrated in FIGS. 6 through 8.

The present invention includes a useful method for identifying peptides that can inhibit the interactions of proteins with other molecules including substrates and other proteins, such as in enzymatic protein-protein

interactions. This can be especially important in modulating activation of proteins involved in complex biological processes such as complement activation. By modulating complement activation, inhibitory peptides may be useful for preventing and treating human pathological conditions associated with activation of the complement system including tissue rejection associated with xenotransplantation of organs, limb and gut ischemia, ischemia-reperfusion following myocardial infarction, stroke, aneurysm, hemorrhagic shock, and crush or thermal injury, anaphylaxis, and any of a variety of chronic inflammation conditions. Autoimmune disorders associated with increased complement activity, including systemic lupus erythmatosis, rheumatoid arthritis and multiple sclerosis may also be regulated by inhibitory peptides or antibodies generated using peptides designed by the indel method. Thus, these new therapeutics may improve the current treatments of these pathological conditions.

For formulations of a pharmaceutical composition containing a biologically active peptide, peptide analog or peptidomimetic compound identified by identifying an indel-associated peptide sequence, an effective amount of the peptide, peptide analog, peptidomimetic or mixtures thereof, is admixed with a physiologically acceptable carrier suitable for administration to mammals including humans. The peptides, peptide analogs or peptidomimetics may be covalently attached to each other, to other peptides or protein carriers or to other carriers such as by incorporation into lipid vesicles. Moreover, for peptides, analogs or peptidomimetics that mediate their biological effect indirectly by eliciting an anti-peptide immune response in the recipient mammal, the peptide, peptide analog or peptidomimetics are mixed with an adjuvant or adsorbent as is well known in vaccine art. The peptides, peptide analogs and/or peptidomimetics may be delivered to the mammal to be treated in any of a variety of known methods including but not limited to systemic delivery via i.v., i.m., i.d. or s.c., or i.p. injection of a solution, suspension or lipid encapsulated form.

Although the examples show use of the method to identify inhibitory peptides, it will be appreciated that the indel method may also be used to look for peptides that potentiate or enhance protein-protein interactions. Because peptides near to or overlapping with an indel sequence represent sections of protein that have a higher potential for being involved in protein-protein interactions, such peptides may also be used to potentiate any reaction dependent on protein-protein interaction. One such example of a complement potentiating peptide is peptide C5-D2 discussed in Example 3. Although the mechanism of potentiation is unknown, one explanation is that the peptide interferes with binding of a complement down-regulatory protein, resulting in increased complement activity.

One general mechanism of peptide-mediated activity is that an indel-proximal peptide selected for a target protein binds to a receptor of that target protein and acts either as an agonist or an antagonist of the normal cellular response associated with protein-receptor binding. For example, a growth factor peptide identified by its indel proximity may instigate or potentiate cell division when the peptide binds to the cellular receptor for the growth factor protein. The indel method is used to select and design peptides, peptide analogs and peptidomemetics that are then tested for their ability to potentiate or enhance cellular responses. Such testing involves the use of well-known *in vivo* or *in vitro* assays available for a wide variety of known cellular responses.

A variety of animal models are available to test the efficacy of peptides identified by the indel method for their relevance to the treatment of human medical conditions. These include *in vivo* animal models include the

following nonlimiting models for: acute myocardial infarction (Weisman. H.F. et al., 1990. *Science* 249: 146); rejection of xenograft transplants (Leventhal, J. et al., 1993, *Transplantation* 55: 857); ischemia related to stroke (Chang, L. et al.. 1992, *J. Cerebr. Blood Flow Metab.* 12: 1030); cardiopulmonary bypass (Nilsson, L. et al., 1990. *Artif. Organs* 14: 46); pancreatitis (Steer, M.. 1992, *Yale J. Biol. Med.* 65: 421) and nephritis (Picler, R. et al., 1994. *Am. J. Pathol.* 144: 915). *Ex vivo* perfusion of animal organs (e.g., heart) may serve as a measurement for peptide inhibition of complement-mediated tissue destruction during xenotransplantation (Morgan, B.P., 1995, *Immunol. Today* 16: 257).

Inhibitory peptides, peptidomimetics, antibodies to peptides or combinations thereof may be administered by a variety of methods well known in the art including intravenous injection, oral and intranasal routes, intraperitoneal, intradermal, intramuscular and subcutaneous injection. A variety of delivery systems may be employed including injection of pharmacologically acceptable solutions or suspensions, encapsulation in liposomes or by controlled release methods well known in the art. *In vitro* treatment of donor organs with prior to transplantation into a recipient is contemplated. Pharmaceutical formulations containing peptides, peptidomimetics, antibodies to peptides or combinations thereof in the range of about 10 μg/kg to 1 g/kg may include other active ingredients including antibiotics, immunosuppressive drugs and growth factors. Animals made transgenic for inhibitory peptides where the peptide is expressed in the animal's organs may also serve as a source of organs for xenotransplantation to avoid hyperacute rejection (Morgan, B.P., 1995, *Immunol. Today* 16: 257).

The disclosed method is a general method for identifying potential interactive sites in a protein and for using this information to guide the design of peptides, peptide analogs, or peptidomimetics that inhibit, enhance, or mimic the activity of a target protein. The method includes the following steps.

First, the known amino acid sequences of two or more proteins that share similar amino acid sequences are compared as described earlier. One of the amino acid sequences is that of the target protein for which modulating/mimicking peptides are sought. The similarity between the protein sequences based on sequences of identical and/or conserved amino acid residues between proteins for comparison should be in the range of about 0.25% to about 70% for identity, about 1% to about 20% for conserved amino acid residues and about 1% to about 80% for a combination of identical and conserved amino acid residues. The preferred ranges are about 3% to about 25% sequence identity and most preferably about 5% to about 15% sequence identity between the compared amino acid sequences. Preferably the compared sequences have about 4% to about 15% conserved amino acid residues and most preferably about 8% to about 12% conserved amino acid residues. The compared sequences have preferably about 7% to about 50% for a combination of identical and conserved amino acid sequence and most preferably about 12% to about 30% for the combined identical and conserved amino acid residues. It will be understood that these numbers represent the average of identical and/or conserved amino acid residues over the entire amino acid sequences of the compared proteins and that limited portions of the sequences may have greater or lesser percentages of identical or conserved amino acid residues.

After aligning the sequences, the indel(s) in the group of proteins are identified. From these identified indels, a peptide sequence that corresponds to a segment of the protein sequence and includes amino acid residues that span or occur within about one to about ten residues of an amino terminus or a carboxyl terminus of an indel are

selected. Peptides are selected using the two guidelines described earlier, and are synthesized using any of a variety of molecular genetic and chemical methods. Using well-known assays, the effects of the peptides on the activity of the target protein or proteins that interact with the target protein are evaluated.

It will be appreciated by one skilled in the art that proteins may share limited similarity when their total amino acid sequences are compared, but may still be related for portions of their amino acid sequences. For such proteins, this method would still be useful for identifying peptides associated with indels in those portions of the proteins that share greater amino acid sequence similarity compared to the dissimilar portions of the proteins. Thus, although the total percentage of identical or conserved amino acid sequence between two or more proteins may be relatively low, alignment of even limited portions of the proteins may identify indels useful for identifying peptides or other molecules that can affect the activities of the proteins. Indels are highly variable in size, ranging from one to about 50 residues in length. The length of the indel is not relevant; the identification of an indel, no matter how short or long, is a key step leading to selection of indel-proximal peptides. To be effective, the selected peptide sequence may range in size from about 4 to 25 amino acid residues in length including peptides of about 4 to 15 or 4 to 20 amino acid residues in length.

A peptide of any length whose amino and carboxy termini are both within about 15, 12 or 10 residues of an indel is included within the scope of this invention if the peptide sequence spans a large indel. For example, hypothetical protein sequences "A" and "B" are shown at an indel (indicated by asterisks) and a peptide from sequence "B" that spans the indel is underlined in sequence "B".

Sequence "A": AAAAAAAAAAAA******************AAAAAAAAAAAA

Sequence "B": BBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBBB

The peptide corresponding to the underlined segment in sequence "B" is 30 residues long and within 6 residues of each end of the indel. Hence a longer peptide is included in the invention because of its proximity to the indel by spanning all of or a portion of the indel. For indels that are particularly long, any peptide that has proximity to the indel, including those completely located within the sequence that corresponds to the indel are within the scope of this invention. For example, using the Sequences "A" and "B" above, a peptide located within the indel would be a series of residues from "B" sequence that corresponds to the sequence indicated by the series of asterisks in the "A" sequence.

Selection of the indel-proximal peptide may be based on peptide characteristics including length, solubility and ease of synthesis or purification, all of which are routine determinations for those skilled in the art of peptide synthesis. The method of delivery of the peptide (e.g., in an aqueous solution or emulsion) may determine the particular indel-proximal peptide sequence selected. Preferably, for any particular indel, an overlapping set of peptides proximal to that indel is selected and each is tested for its ability to alter the activity of the target protein in a variety of assays. Thus, an effective indel-proximal peptide may span the indel site or flank the indel site. Peptides that flank an indel site may occur proximal to either the amino or carboxyl terminus of the indel and may have at least one residue within about one, five, ten, fifteen, twenty or thirty amino acid residues of one terminus of the indel.

Because this is a general method useful for identifying peptides that affect or mimic protein-protein interactions, it is useful for any family of proteins for which amino acid sequences are known. The amino acid sequences may be known for an entire protein, or only a portion of a protein, within a family of proteins and may be predicted based on the nucleic acid sequence coding for a protein. Thus for any family of proteins having amino acid sequences known for at least two members of the family, the method is useful for identifying peptides that modulate or mimic protein activities.

The general principles of the present invention may be more fully appreciated by reference to the following nonlimiting examples which demonstrate the utility of the method for identifying inhibitory peptides for the complement protein C3/C4/C5 and C2/B families.

EXAMPLE 1: Identification of indel-proximal sites can be used to design specific peptide inhibitors.

An alignment of the C3/C4/C5 family of complement proteins using the CLUSTAL W program revealed over 30 indels as shown in FIG. 2A-2E. Indels were somewhat arbitrarily defined here as loci where one member of the family had an insertion or deletion equal to or greater than two residues in length. However, it will be understood that other definitions are also predictive of inhibitory peptides based on this general procedure of identification, including where one member of the family has an insertion or deletion equal to or greater than one amino acid residue in length. In FIG. 2A-2E, the indels are identified by a solid line under the last line of protein sequence for each section of six sequences and labelled with consecutive numbers. The amino acid sequences of the peptides that were synthesized on the basis of their proximity to these indels are underlined in the human C3, C4 or C5 amino acid sequences of FIG. 2A-2E and listed in Table 3 with their corresponding SEQ ID NO.

Previously reported interactive sites in C3, C4, and C5 (shown as lower case letters for the amino acid residues in FIG. 2A-2E) were also found to cluster around the indels (e.g., see those listed in Table 1). These putative interactive sites have been identified previously by other methods as binding sites to complement control proteins, receptors, associating subunits in the complex convertases, and numerous cleavage sites by activating and inactivating proteases.

Only a few notable interactive sites were not near indels. These include the thioester site and the proteolytic site that occurs eight residues upstream of the thioester in the C3 α-chain that generates the C3d fragment from C3dg, which may not be physiologically important (Law, S.K.A., 1988, *J. Cell Sci. Suppl.* 9:67-97; DeBruijn, M.H.L. & Fey, G.H., 1985, *Proc. Natl. Acad. Sci. USA* 82:708-712). Another is the isotype-specific site in C4, which may be involved in *intramolecular* interactions (Law, S.K.A. et al., 1984, *EMBO J.* 3:1819-1823; Isenman, D.E. & Young, J.R., 1984, *J. Immunol.* 132:3019-3027; Yu, C.Y. et al., 1986, *EMBO J.* 5: 2873-2881). Thus, the majority of active sites in the C3, C4 and C5 complement family were proximal to indels identified by the present method.

For use in the following examples, enzymes and antisera can be purified using well known techniques. Most complement reagents are commercially available (e.g., from Advanced Research Technologies, San Diego, CA). Other commercial sources include: human C1s (Enzyme research Labs. Inc., South Bend, IN), human factor B (Calbiochem, La Jolla, CA), cobra venom factors (Diamedix, Miami, FL or Quidel, San Diego, CA), human factor D, (Quidel, San

Diego, CA), and goat antisera against murine C3 (Organon Teknika-Capel, West Chester, PA) and human C3 and C5 (Quidel, San Diego, CA). Synthetic peptides may be obtained from Chiron Mimotopes (San Diego, CA).

Table 3. Inhibitory activity of C3/C4/C5 peptides in hemolytic assays.

SEQ ID NO:	Name	Indel[1]	Sequence[2]	Hydrop[3]	IH_{50}[4]
54	III-11	24	KAFSDRNTLIIYLD-NH_2	0.44	25
55	II-4	11	Ac-EVVADSVWVDVKDS-NH_2	0.33	50
53	III-8C	23	Ac-SILDISMMTGFAPDT-NH_2	0.63	60
56	II-5	13/14	Ac-SEFPESWLWNVEDL-NH_2	0.59	100
57	A15	27/28	Ac-LSSDFWGEKPNLS	0.39	200
58	A12	9	VNFLLRMDRAHEAK-NH_2	0.24	200
59	II-1	1	Ac-AQGDVPVTVTVHD-NH_2	0.37	250
60	A14	11	Ac-SGQREVVADSVWVDV	0.34	250
61	I-1	2	TIPANREFKSEKGR	-0.07	300
62	I-4[5]	8	SITVRTKKQELSEA	0.07	300
63	I-8	14	DLKEPPKNGISTKL	0.13	300
64	II-3	7	Ac-GDGVAKLSINTHPS-NH_2	0.26	>300
51	II-8A[5]	23	Ac-DATMSILDISMMTG-NH_2	0.59	>300
65	I-11	16	Ac-ERLGREGVQKEDI	-0.09	375
66	I-6	11	YYTLIGASGQREVV	0.47	400
67	III-12	7	Ac-DGSPAYRVPVAVQGE-NH_2	0.27	400
52	III-8B	23	Ac-DQDATMSILDISMM-NH_2	0.50	>400
68	I-5	10	RLLKAGRQVREPGQ	0.04	450
69	II-6	15	PKSSLSVPYVIVP-NH_2	0.71	>600
70	II-7	20	Ac-QVNSLPGSITKAGD-NH_2	0.24	>600
71	III-1	2	TVLTPATNHMGNVT-NH_2	0.45	>600
72	C4-B1	2	Ac-EVQLVAHSPWLKDS	0.47	150
73	C5-D5	7	Ac-TSDLDPSKSVTRVD	0.07	200
74	C5-D2	11	Ac-TAELVSDSVWLNIE	0.55	•

1 Indel(s) proximal to the peptide sequence in the intact protein.

2 Unless otherwise indicated by "C4" or "C5" in the peptide name, the peptides have C3 protein sequences. Acetylated (Ac-) amino termini and amidated (-NH_2) carboxyl termini are as noted.

3 Hydropathy of the peptide indicated using a scale where values about 0 indicate very hydrophilic, and values about 1 indicate very hydrophobic. Hydropathies were calculated as the mean of the amino acid parameters described by Fauchere, J.L. et al., 1983, *Eur. J. Med. Chem.* 10:369.

4 Peptide concentration (μM) giving 50% inhibition of hemolytic activity in about 0.15% human serum, which gives about 30% hemolysis of input target EA in the absence of peptide.

5 Peptide alone causes hemolysis.

• Enhances hemolytic activity 3-fold, with 50% effect (2-fold increase) at about 80 μM.

EXAMPLE 2: Synthetic Peptides Predicted by Indel Analysis

The peptides listed in Table 3, representing indel-proximal sequences in the C3, C4 and C5 proteins, were tested for inhibition of complement function. The peptides were synthesized using well-known methods (Geysen, H.M. et al., 1984, *Proc. Natl. Acad. Sci. USA* 81:3998; Geysen, H.M. et al., 1983, *J. Immunol. Meth.* 102: 259).

This group of indel-proximal peptides for the C3, C4 and C5 family of proteins are underlined in FIG. 2A-2E. The names of the individual peptides (e.g., II-1) used in Table 3 correspond to the names over the underlined peptide sequences in FIG. 2A-2E. Some of the peptides were acetylated at the amino terminus (indicated by "Ac-" in Table 3) or were amidated at the carboxyl terminus (indicated by "-NH_2 in Table 3). The results presented in column 6 of Table 3 are discussed in detail below at Example 3.

Although these peptides were synthesized using specific chemical methods, it will be appreciated by those skilled in the art that other methods of producing the peptides would be equivalent. That is, other chemical procedures may be used to make peptides. Moreover, peptides may be produced by degradation of larger proteins containing the peptide sequence followed by purification of the peptide by well known methods. Recombinant genetic techniques may be used to express a DNA sequence coding for a peptide in either microbial or mammalian cells followed by purification of the peptide. Therefore, peptides selected on the basis of the general indel method could be produced by a variety of procedures well known to those skilled in the art and the invention should not be considered limited to any one particular method of making the peptides so identified.

All of the synthesized peptides were tested for their ability to inhibit *in vitro* hemolysis of sensitized erythrocytes by human complement.

EXAMPLE 3: Hemolysis of Erythrocytes

Peptides identified in Example 2 were tested for inhibition of complement activity using the hemolytic assay for complement activity. The peptides were added to the hemolytic assay and their ability to inhibit lysis of erythrocytes or red blood cells (RBC) by complement activity was measured.

In this assay, sheep RBC coated with rabbit anti-sheep RBC antibodies were exposed to dilute human serum. Complement is activated on erythrocyte surfaces by the rabbit antibody resulting in lysis of the red blood cells. The amount of complement in the serum is determined by the extent of erythrocyte lysis. If peptides inhibit any step in the lytic pathway, this inhibition reduces the apparent complement activity of the serum resulting in less erythrocyte lysis.

Erythrocyte lysis or hemolytic assays were performed using the "EZ Complement" kit (Diamedix, Miami, FL). To 0.5 ml aliquots of sensitized erythrocytes were added 3 μl of a 1:4 dilution of a standard human serum sample, a control human serum sample, the peptide of interest, and the control sample plus the peptide. The latter two samples measured the hemolytic activity of the peptide alone and the inhibitory effect of the peptide on complement lysis, respectively. The level of hemolysis was measured as the absorbance of the solution at 415 nm after centrifugation of the sample to remove intact erythrocytes. The level of hemolysis was expressed as the fraction

of hemolysis of the control sample alone, and subtracting the hemolysis induced by the peptide alone, without added control.

FIG. 9 shows the hemolysis inhibition results for C3 peptides II-4, II-5, III-8C and III-11. The lytic activity of the peptide alone was negligible for these peptides. The data for the peptides (II-4, II-5, III-8C and III-11) shown in FIG. 9 show that increasing concentrations of peptide resulted in decreasing complement activity. The concentration of peptide necessary for 50% inhibition differed for each peptide, but was as low as about 25 μM for peptide III-11. Therefore, the hemolysis data indicated that some of the C3 peptides inhibited complement function at concentrations consistent with those seen for other inhibitory interface peptides (Dutia, B.M. et al., 1986, *Nature* 321: 439-441; Haigh, A. et al., 1990, *Nature* 344: 257-259; Babe, L.M. et al., 1992, *Protein Sci.* 1: 1244-1253).

The results in FIG. 9 were selected to show the range of responses found among the most strongly inhibiting peptides. Results for all the peptides are listed in Table 3, which presents the approximate peptide concentration at which 50% inhibition of erythrocyte lysis (IH_{50}) was seen. These results are shown graphically in FIG. 10.

The hemolytic results suggest that specific peptide-protein interactions interfere with the specific protein-protein interactions that are necessary for complement function. Inhibition of hemolysis was dose-dependent, and the concentration of peptide necessary for 50% inhibition varied substantially with each peptide (from about 25 μM to greater than about 600 μM). These results show that some peptides inhibit the complement system specifically. Moreover, the range of peptide concentrations needed for inhibition is comparable to that seen in studies of other inhibitory peptides in which the concentration of peptide necessary for inhibiting a protein-protein interaction was about 100 μM. This result is especially true for inhibition by peptides III-11, II-4, III-8C, and II-5.

The mechanism(s) by which the peptides inhibit hemolytic activity of complement is unclear. Inhibition by C3 peptides may result from inhibition of C3 interaction with one or several other proteins of the complement cascade such as the interaction of substrate C3 with C4b or C2a in the classical pathway convertase. That is, peptides may block formation of the complex with C4b and C2a after cleavage by the C3 convertase, and/or they may block recognition of C5 by the C4bC3bC2a complex. Any of these effects results in loss of convertase function and inhibition of the complement pathway.

Table 3 includes data for a single human C4 peptide (peptide C4-B1) and two human C5 peptides (peptides C5-D5 and C5-D2, respectively), that were also tested for inhibition of hemolytic activity. A total of seven C5 peptides have been examined to date. Peptide C4-B1, the only C4 peptide chosen for study, shows moderate inhibitory activity. Thus, in a single attempt, the indel-proximal method of peptide identification successfully located an inhibitory peptide in a second member of the C3/C4/C5 family. Peptide C5-D5 shows the greatest inhibitory activity of the C5 peptides tested to date although its activity is relatively weak. In contrast, peptide C5-D2 shows strong potentiating activity in the hemolytic assay, possibly due to inhibition of C5 interaction with a natural inhibitor protein in human serum. Based on the results obtained with peptides C5-D5 and C5-D2, the disclosed method can be used to identify peptides that have potentiating as well as attenuating effects in a complex biochemical system.

The inhibitory peptides did not show any pattern of favored orientation with regard to indels. That is, inhibitory peptides spanned and flanked both the N- and C- termini of indels and came from indels where there were deletions and insertions in the C3 sequence relative to the C4 and/or C5 sequences. Moreover, there was no preferred amino acid composition. The peptide length chosen arbitrarily in these examples was 13 to 15 amino acid residues.

In the next example, C2 peptides identified by the indel method were synthesized and tested for inhibition of complement activity as shown by the data in Table 4.

Table 4. Inhibitory activity of C2 peptides in hemolytic assays.

SEQ ID NO:	Name	Sequence[1]	Hydrop[2]	IH_{50}[3]
75	E1	Ac-MRLLGMETMAWQE	0.57	30
76	E2	REILNINQKRNDY-NH_2	-0.01	>300
77	E4	WRVNVGDPKSQWGK-NH_2	0.22	>300

1 The single-letter amino acid code is used with Acetylated (Ac-) amino termini and amidated (-NH_2) carboxy termini.

2 Hydropathy of the peptide indicated using a scale where values about 0 indicate very hydrophilic, and values about 1 indicate very hydrophobic. Hydropathies were calculated as the mean of the amino acid parameters described by Fauchere, J.L. et al., 1983, *Eur. J. Med. Chem.* 10:369.

3 Peptide concentration (μM) giving 50% inhibition of hemolytic activity (CH_{50}) in about 0.15% human serum, which gives about 30% hemolysis of input target EA in the absence of peptide.

EXAMPLE 4: Peptide inhibitors from the C2/factorB complement protein family designed by indel association.

To test further the effectiveness of the indel-proximal method, the C2/factor B family of complement proteins was analyzed. C2 and factor B are structurally similar proteins (about 100,000 m.w.) that play analogous roles in the classical and alternative pathways, respectively, of complement activation and whose genes arose from a common ancestor (Campbell, R.D. et al., 1988, *Ann. Rev. Immunol.* 6:161-195). Using procedures essentially as described in Example 1, members of the C2/factor B complement protein family were aligned as shown in FIG. 11A-11B, indels were identified and three indel-proximal peptides were selected for testing (underlined in FIG. 11A-11B). As seen by the data in Table 4, two peptides, E2 and E4 (SEQ ID NO:76 and SEQ ID NO:77, respectively), showed no evidence of inhibitory activity in the hemolytic assay. Peptide E1 (SEQ ID NO:75), however, showed very potent activity, as shown in FIG. 12 which presents inhibition of complement hemolytic activity by peptide E1 in the presence and absence of serum in the assay. About 30 μM of peptide E1 resulted in 50% inhibition, which is essentially equal to the most potent C3 inhibitor, peptide III-11 (SEQ ID NO:54).

Hence the indel-proximal method of peptide identification has been successfully used to guide the design of peptide inhibitors for another protein family in the complement system by efficiently identifying protein interactive regions and inhibitory peptides which are presumably interface peptides.

EXAMPLE 5: Generation of Antibodies Against Indel-associated Peptides

Antibodies that recognize the inhibitory peptides identified by the indel method may also be useful reagents for inhibiting protein activity. Such antibodies would be expected to bind to the indel-proximal portion of the native protein, thus blocking activity associated with that portion of the protein.

Polyclonal and monoclonal antibodies can be generated against the purified peptides designed by using the indel method. Monoclonal antibodies useful in the present invention can be produced and isolated by processes which are well known in the art, such as those discussed by Milstein and Kohler (*Nature* 256:495-497, 1975). Monoclonal antibodies may be more useful because of the specificity associated with such antibodies and the ability to produce antibodies with the same specificity from a clonal cell line. Peptides designed by using the indel method that have inhibitory activity in an *in vitro* or *in vivo* assay are synthesized and monoclonal antibodies are then generated using the purified peptides by standard procedures.

In accordance with one known process for preparing monoclonal antibodies, mice such as Balb/c female mice or other mouse strains or even other suitable animals (e.g., rats or rabbits) are immunized with an amount of a peptide of interest, such as those identified in Examples 3 and 4, to initiate an immune response. The animals are immunized with the peptides mixed with a suitable adjuvant or with peptides conjugated to a carrier molecule. The peptide dosage and immunization schedule for producing useful quantities of suitable splenocytes can be readily determined by one skilled in the art depending on the animal strain used.

The size and spacing of doses of peptide are generally microgram quantities with a minimum dosage for initiating an immune response typically in the range of 0.1-100 μg/animal. For example, an initial immunization with approximately 50 μg of peptide may be followed by a series of five injections for hyperimmunization. An adjuvant (e.g., Freund's incomplete and complete adjuvants and alum gels) may be mixed with the peptide antigen to enhance

antibody production against the antigen using methods well known in the art. Thus, a given dose of peptide may be more effective when injected subcutaneously with an adjuvant or when injected as repeated small aliquots than when administered intravenously.

Following the primary immunization with the peptide antigen, the animal is monitored for production of anti-peptide specific serum antibodies using well known techniques (e.g., ELISA or radioimmunoassay) normally about one to two weeks after immunization. After the primary response to the peptide antigen is detected, a second dose of the same antigen is given to elicit a peptide-specific secondary immune response which is also detected by standard immunoassays. After one to five booster immunizations, the animal is killed and its spleen cells are isolated and fused with myeloma cells (e.g., the murine cell line Sp2/0-Ag14) to produce hybridoma cell lines capable of reproduction *in vitro* to produce anti-peptide antibodies.

The myeloma cell line selected should be compatible with the spleen cells, and optimally should be a cell line of the same species as the spleen cells. Although the murine cell line Sp2/0-Ag14 has been found to be effective for use with mouse spleen cells, other myeloma cell lines can alternatively be used. See, for example, Nature, 276: 269-270 (1978) and U.S. Patent No. 5,472,868, for fusion partner cells and methods of using such cells to produce hybridomas.

Spleen cells are fused with an appropriate myeloma cell line using polyethylene glycol. After fusion of spleen and myeloma cells, the mixture of unfused spleen cells, unfused myeloma cells and fused cells are diluted and cultured in a selective medium (e.g., containing hypoxanthine, aminopterin and thymidine) that supports growth of fused cells and will not support the growth of the unfused myeloma cells for a time sufficient to allow death of all unfused cells. Since the unfused spleen cells are nonmalignant, they have only a finite number of generations until they fail to reproduce. The fused cells reproduce because they possess the malignant quality contributed by the myeloma parent and the enzyme necessary to survive in the selected medium contributed by the spleen cell parent.

The supernatant from each of a plurality of hybridoma-containing tissue culture wells is evaluated for the presence of antibody specific to the peptide of interest using any of a number of well known immunoassays. Hybridomas that produce antibodies that specifically recognize the peptide used as an immunogen are cloned (e.g., by limiting dilution) for subsequent production of antibodies *in vitro* or *in vivo*. Anti-peptide antibodies are produced either *in vitro* by tissue culture of the selected cell lines or *in vivo* by generating ascites fluid in mice injected with the hybridoma cell line. The monoclonal antibody may then be isolated in accordance with techniques known in the art.

Purified Anti-peptide antibodies are identified as those capable of modulating a biological system, such as the complement system, using the appropriate *in vitro* or *in vivo* assay for the target protein or protein system. For example, the monoclonal antibodies are tested for their ability to enhance or inhibit hemolysis in the assay as described in Example 3. The anti-peptide antibody is added into the hemolysis assay alone or in combination with the corresponding peptide against which the antibody was raised. To further test antibodies identified as having a biological activity, they are tested for their ability to mediate a therapeutic effect as in the *in vivo* assay described in Example 6. That is, the anti-peptide antibody is injected before, simultaneous or after injection of the

corresponding peptide or instead of the peptide during the antibody challenge and modulation of the inflammation response is monitored relative to the appropriate control reactions.

EXAMPLE 6: *In Vivo* Assay to Measure Peptide Inhibition of Complement-mediated Inflammation

The Arthus model of complement-mediated inflammation is caused by interaction of tissue antigen with circulating antibody. This *in vivo* model of complement activation is characterized by formation of immune complexes resulting in complement activation, inflammatory cell recruitment, edema and tissue damage (Bailey, P. & Sturm, A., 1983, *Biochem. Pharm.* 32: 475). A passive Arthus reaction is established in an animal (e.g., a guinea pig) by first injecting an antigen i.v. and then challenging with an antigen-specific antibody. Peptide inhibition of the complement-mediated response is measured when peptide is injected i.d. before or simultaneous with antibody challenge and then biopsy tissue taken from the antibody challenge site is assayed for inflammation.

Male guinea pigs (about 300 g) are anesthetized by injection of sodium pentobarbital (40 mg/kg, i.p.) and then injected i.v. with ovalbumin (20 mg/kg) and ^{125}I-labeled bovine serum albumin (^{125}I-BSA, 1 μCi). Antibody challenge immediately follows by injecting animals in the dorsal region with polyclonal anti-ovalbumin antibody (10 mg, i.d.), with or without an indel identified peptide in the μM range indicated as effective complement inhibition by the hemolysis assay (see Example 3).

Three hours after injection of the antibody challenge, the animals are humanely killed and skin tissue from the antibody challenge site is excised by biopsy punch (about 2-5 mm). Inflammation is measured by leakage of ^{125}I-BSA into the skin tissue determined by standard radiochemical counting procedures for ^{125}I; percent inhibition by peptide is measured by comparing leakage of ^{125}I-BSA into the skin tissue in the presence of peptide with that seen in the absence of peptide. As a positive control for complement inhibition, a separate animal is injected with cobra venom factor (200 U/kg, i.p.) 24 hr before initiation of the Arthus response; the venom factor results in suppression of the complement response in animals that receive antigen and antibody without complement-inhibitory peptide.

Peptides identified as being complement inhibitory in the hemolysis assay show complement inhibition in this *in vivo* model in a dose-dependent manner with maximal inhibition comparable to the animal that received cobra venom factor before initiation of the Arthus response. For example, a series of 250-300 g male guinea pigs are anesthetized (with sodium pentobarbital, at 40 mg/kg, i.p.) and then injected i.v. with ovalbumin (20 mg/kg) and ^{125}I-labeled bovine serum albumin (^{125}I-BSA, 1 μCi). One of the animals is the cobra venom factor control which received cobra venom factor (200 U/kg, i.p.) at about 24 hr before injection of the ovalbumin and ^{125}I-BSA. The other animals are individually i.v. injected with 0.1 μg, 0.5 μg, 1 μg, 5 μg and 10 μg of peptide III-8C (SEQ ID NO:53), or no peptide, simultaneously with the ovalbumin and ^{125}I-BSA i.v. injection. Each animal then receives polyclonal anti-ovalbumin antibody (10 mg, i.d. in the dorsal region). Three hours after injection of the anti-ovalbumin antibody challenge, the animals are humanely killed and skin tissue from the antibody challenge site is excised (by biopsy punch, 2-5 mm). Inflammation is measured by leakage of ^{125}I-BSA into the skin tissue determined by standard radiochemical counting procedures for ^{125}I to determine the percentage of inhibition of inflammation mediated by the peptide III-8C dosage. The inhibition seen in the venom factor control animal serves as an arbitrary 100% inhibition data point whereas the data for the animal that received no peptide and no venom factor serves as an arbitrary 0%

inhibition data point. By comparison with the results in control animals, the animals that receive peptide III-8C show about 5-10% inhibition with 0.1 μg peptide, about 7-21% inhibition with 0.5 μg, about 10-35% inhibition with 1 μg, about 20-45% inhibition with 5 μg, and about 40-80% inhibition with 10 μg of peptide III-8C. Thus, *in vivo* activity of a peptide parallels that seen *in vitro*. Although the mechanism of this activity is not known, it may result from inhibition of one or more proteins of the complement system, competition with C3 protein, or binding to a protein or ligand such as a C3 convertase.

EXAMPLE 7: Selection and/or Synthesis of Peptidomimetic Inhibitors Based on Indel-identified Peptides

Because peptidomimetic molecules can exhibit increased potency as inhibitory molecules for protein-protein interactions, peptidomimetics of an inhibitory peptide identified using the indel method and the hemolytic assay are selected and screened for increased inhibitory activity.

Peptidomimetics of peptide II-1 (SEQ ID NO:59) are selected from a combinatorial peptide library in which the internal five amino acid residues (VPVTV) are randomized (Bianchi, E. et al., 1995, *J. Mol. Biol.* 247: 154-160) producing a library of peptides that are analogs of peptide II-1. The library of peptides can be produced either using molecular genetic techniques (e.g., synthesis of the partially randomized DNA sequences coding for the peptide II-1 analogs followed by expression of the peptides in cells and purification of the peptides) or by chemically synthesizing the partially randomized peptides. Chemical synthesis allows additional expansion of the repertoire by allowing inclusion of non-coded amino acids or organic isosteric replacement groups (e.g., a 4(5)-acylimidazole ring for replacement of an amide bond). The peptide analogs and peptidomimetics are tested *in vitro* for their inhibitory activity relative to peptide II-1 in a hemolytic assay essentially as described in Examples 3 and 4 and in an *in vivo* assay essentially as described in Example 6. Some of the partially randomized peptidomimetics and analogs of peptide II-1 show increased inhibition of complement function in the hemolysis assay, some show increased inhibition the animal model and some show increased inhibitory activity in both assays.

Peptide I-8 (SEQ ID NO:63) is used to synthesize a retro-D-amino acid peptide that is an isomer of the corresponding L-amino acid I-8 peptide. That is, a peptide is synthesized using only D-amino acids and having a sequence LKTSIGNKPPEKLD (SEQ ID NO:78) which is the reverse of the sequence of peptide I-8. The I-8 peptide and the retro-D-peptide of I-8 are compared in the hemolytic assay as described in Example 3. The I-8 shows little or no inhibition of hemolysis, requiring a peptide concentration of about 300 μM to produce 50% inhibition of the hemolytic activity in 0.15% human serum. In contrast, the retro-D-peptide of I-8 having reversed peptide bond orientation of the corresponding L-amino acid peptide shows increased inhibition because it produces 50% inhibition at a concentration of about 150 μM in the same type of reaction.

Another peptide analog is produced as a covalently cyclized version of peptide I-5 having the linear sequence CRLLKAGRQVREPGQC (SEQ ID NO:79) (i.e., addition of cysteine residues at the amino- and carboxyl-termini of peptide I-5). The linear peptide having SEQ ID NO:79 is synthesized using conventional solid-phase protein synthesis and purified by reverse phase HPLC using standard methods. The purified peptide is then oxidized to covalently cyclize the terminal cysteine sulphydryl side chains to form an intrachain disulfide bridge (such as described by McDonnell, J.M. et al., 1996, *Nature Structural Biol.* 3(5):419-426), thus constraining the peptide structurally. Moreover, the

corresponding linear version of the peptide is used as a control and all three peptides (I-5, the linear and cyclized forms of SEQ ID NO:79) are compared in the hemolytic assay as described in Example 3. The I-5 peptide shows limited inhibition of hemolysis, requiring a peptide concentration of about 450 μM to produce 50% inhibition of the hemolytic activity in 0.15% human serum. Similarly, the linear version of the peptide having SEQ ID NO:79 also shows limited inhibition of hemolysis, requiring a peptide concentration of about 350-400 μM to produce 50% inhibition of the hemolytic activity in 0.15% human serum. In contrast, the cyclized version of the peptide having SEQ ID NO:79 produces no inhibition of hemolysis and enhances hemolytic activity about 2-fold at a concentration of about 100 μM in the same type of reaction. Thus the physical form of the peptide having SEQ ID NO:79 substantially affects its ability to modulate protein activity in this assay.

EXAMPLE 8: Indel-associated peptide substitution for a target protein's receptor

A series of peptides are generated for the sequence of complement C3 protein near indel 13 as shown in FIG. 2A-2E, such as peptides II-5 (SEQ ID NO:56) and I-8 (SEQ ID NO:63), and cyclized peptidomimetics of these peptides in which the amino- and carboxyl-termini D-cysteines joined by a disulfide bridge are synthesized using techniques described earlier. Peptide I-11 (SEQ ID NO:65), overlapping indel 16 in FIG. 2A-2E, and its corresponding cyclized-D-cysteine peptidomimetic is used as a negative control and purified C3 protein is used as a positive control. These peptides and protein are radiolabeled using standard procedures and the labeled peptides and protein are tested for their ability to bind to complement receptors CR1, CR2 and CR3 using *in vitro* binding assays known in the art (Diefenbach, R.J. & Isenman, D.E., 1995, *J. Immunol.* 154:2303; Taniguchi-Sidle, A. & Isenman, D.E., 1994, *J. Immunol.* 153:5285). These binding assays show that purified C3 binds to complement receptors CR1, CR2 and CR3, neither the linear peptide I-11 or its cyclized peptidomimetic bind to any of the receptors, and the linear forms of peptides II-5 and I-8 bind to all three C3 receptors but never more than about 40-50% of the efficiency of binding as with purified C3 protein. The cyclized peptidomimetic of peptide II-5 binds to all three receptors but less efficiently than its linear form whereas the cyclized peptidomimetic of peptide I-8 binds slightly more efficiently to all three receptors than the linear peptide. These results show that peptides identified by the indel-association method and peptidomimetics modeled on such peptides can substitute for a ligand or substrate of the target protein.

EXAMPLE 9: Pharmaceutical composition including an indel-proximal identified peptide or peptide analog or peptidomimetic thereof

Peptides, peptide analogs or peptidomimetic molecules according to the present invention are identified, synthesized and tested for biological activity in an *in vitro* and/or *in vivo* assay as described in the previous examples or using other well known methods and assays. For delivery of a peptide, peptide analog or peptidomimetic molecule (hereafter generally referred to as "peptide"), the peptide is dissolved in a buffered salt solution (e.g., phosphate buffered saline) at a dose of 1 μg/ml to 100 μg/ml depending on the activity of the peptide as determined in an *in vivo* or *in vitro* assay. The peptide solution is then injected in 0.5-5 ml aliquots into the mammal including a human to be treated via any of a variety well known routes (i.v., i.m., s.c., etc.) depending on the type of treatment. For example, for a localized inflammatory response, a peptide that inhibits complement activity (e.g., peptide III-11) is injected locally near the area of inflammation, with repeated dosages periodically until inflammation is sufficiently

decreased as determined by the treating physician. In another application, a patient experiencing ischemia-reperfusion following myocardial infarction may be treated with a complement-inhibiting peptide delivered by catheterization to deliver the peptide selectively to the affected area.

For peptides that have been identified as being active by eliciting antibodies to the peptide which then mediate the desired anti-target protein activity, the peptide is preferably coupled to ovalbumin grade V at about a 10:1 molar ratio using N-succinimidyl 3-(2-pyridyldithio) propionate (SPDP). Briefly, SPDP (40 mM in 99% ethanol is added dropwise to ovalbumin (dissolved in 0.2M NaH_2PO_4, pH 8.5) to form a SPDP-ovalbumin conjugate that is purified by column chromatography and coupled to peptide (1 mg/ml in 10% acetic acid) by allowing the peptide and SPDP-ovalbumin to incubate for about 12 hr at RT. The peptide-SPDP-ovalbumin conjugate (100 μg/0.5 ml of phosphate buffered saline, with or without an adjuvant) is injected i.m. in a series of three injections about one week apart to induce an anti-peptide immune response in the mammal receiving the peptide.

Although the present invention has been described in the context of particular examples and preferred embodiments, it will be understood that the invention is not limited to such embodiments. Instead, the scope of the present invention shall be measured by the claims that follow and all modifications which come within the meaning and range of the lawful equivalency of the claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Lidak Pharmaceuticals

(ii) TITLE OF THE INVENTION: METHOD FOR IDENTIFYING PEPTIDES THAT AFFECT PROTEIN-PROTEIN INTERACTIONS AND COMPLEMENT-MODULATING PEPTIDES

(iii) NUMBER OF SEQUENCES: 79

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Knobbe, Martens, Olson and Bear
(B) STREET: 620 Newport Center Blvd. 16th Floor
(C) CITY: Newport Beach
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 92660

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Israelsen, Ned A
(B) REGISTRATION NUMBER: 29,655
(C) REFERENCE/DOCKET NUMBER: LIDAK.048A

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 619-235-8550
(B) TELEFAX: 619-235-0176
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Val Pro Val Ala Val Gln Gly Glu Asp Thr Val Gln Ser Leu Thr Gln
 1               5                  10                  15
Gly Asp Gly
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Val Leu Val Val Thr Gln Gly Ser Asn Ala Lys Ala Leu Thr Gln Asp
 1               5                  10                  15
Asp Gly
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
Ile Pro Val Lys Val Ser Ala Thr Val Ser Ser Pro Gly Ser Val Pro
 1               5                  10                  15
Glu Ala Gln Asp Ile Gln Gln Asn Thr Asp Gly
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Val Pro Val Lys Val Ser Ala Thr Leu Val Ser Gly Ser Asp Ser Gln
 1               5                  10                  15
Val Leu Asp Ile Gln Gln Ser Thr Asn Gly
            20                  25
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Val Pro Val Ile Leu Asn Ala Gln Thr Ile Asp Val Asn Gln Glu Thr
 1               5                  10                  15
Ser Asp Leu Asp Pro Ser Lys Ser Val Thr Arg Val Asp Asp Gly
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
Val Pro Val Thr Leu Met Ala Gln Thr Val Asp Val Asn Gln Glu Thr
 1               5                  10                  15
Ser Asp Leu Glu Thr Lys Arg Ser Ile Thr His Asp Thr Asp Gly
            20                  25                  30
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1663 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
Met Gly Pro Thr Ser Gly Pro Ser Leu Leu Leu Leu Leu Leu Thr His
 1               5                  10                  15
Leu Pro Leu Ala Leu Gly Ser Pro Met Tyr Ser Ile Ile Thr Pro Asn
            20                  25                  30
Ile Leu Arg Leu Glu Ser Glu Glu Thr Met Val Leu Glu Ala His Asp
        35                  40                  45
Ala Gln Gly Asp Val Pro Val Thr Val Thr Val His Asp Phe Pro Gly
    50                  55                  60
Lys Lys Leu Val Leu Ser Ser Glu Lys Thr Val Leu Thr Pro Ala Thr
65                  70                  75                  80
Asn His Met Gly Asn Val Thr Phe Thr Ile Pro Ala Asn Arg Glu Phe
                85                  90                  95
Lys Ser Glu Lys Gly Arg Asn Lys Phe Val Thr Val Gln Ala Thr Phe
            100                 105                 110
Gly Thr Gln Val Val Glu Lys Val Val Leu Val Ser Leu Gln Ser Gly
        115                 120                 125
Tyr Leu Phe Ile Gln Thr Asp Lys Thr Ile Tyr Thr Pro Gly Ser Thr
    130                 135                 140
Val Leu Tyr Arg Ile Phe Thr Val Asn His Lys Leu Leu Pro Val Gly
145                 150                 155                 160
```

Arg Thr Val Met Val Asn Ile Glu Asn Pro Glu Gly Ile Pro Val Lys
165 170 175
Gln Asp Ser Leu Ser Ser Gln Asn Gln Leu Gly Val Leu Pro Leu Ser
180 185 190
Trp Asp Ile Pro Glu Leu Val Asn Met Gly Gln Trp Lys Ile Arg Ala
195 200 205
Tyr Tyr Glu Asn Ser Pro Gln Gln Val Phe Ser Thr Glu Phe Glu Val
210 215 220
Lys Glu Tyr Val Leu Pro Ser Phe Glu Val Ile Val Glu Pro Thr Glu
225 230 235 240
Lys Phe Tyr Tyr Ile Tyr Asn Glu Lys Gly Leu Glu Val Thr Ile Thr
245 250 255
Ala Arg Phe Leu Tyr Gly Lys Lys Val Glu Gly Thr Ala Phe Val Ile
260 265 270
Phe Gly Ile Gln Asp Gly Glu Gln Arg Ile Ser Leu Pro Glu Ser Leu
275 280 285
Lys Arg Ile Pro Ile Glu Asp Gly Ser Gly Glu Val Val Leu Ser Arg
290 295 300
Lys Val Leu Leu Asp Gly Val Gln Asn Leu Arg Ala Glu Asp Leu Val
305 310 315 320
Gly Lys Ser Leu Tyr Val Ser Ala Thr Val Ile Leu His Ser Gly Ser
325 330 335
Asp Met Val Gln Ala Glu Arg Ser Gly Ile Pro Ile Val Thr Ser Pro
340 345 350
Tyr Gln Ile His Phe Thr Lys Thr Pro Lys Tyr Phe Lys Pro Gly Met
355 360 365
Pro Phe Asp Leu Met Val Phe Val Thr Asn Pro Asp Gly Ser Pro Ala
370 375 380
Tyr Arg Val Pro Val Ala Val Gln Gly Glu Asp Thr Val Gln Ser Leu
385 390 395 400
Thr Gln Gly Asp Gly Val Ala Lys Leu Ser Ile Asn Thr His Pro Ser
405 410 415
Gln Lys Pro Leu Ser Ile Thr Val Arg Thr Lys Lys Gln Glu Leu Ser
420 425 430
Glu Ala Glu Gln Ala Thr Arg Thr Met Gln Ala Leu Pro Tyr Ser Thr
435 440 445
Val Gly Asn Ser Asn Asn Tyr Leu His Leu Ser Val Leu Arg Thr Glu
450 455 460
Leu Arg Pro Gly Glu Thr Leu Asn Val Asn Phe Leu Leu Arg Met Asp
465 470 475 480
Arg Ala His Glu Ala Lys Ile Arg Tyr Tyr Thr Tyr Leu Ile Met Asn
485 490 495
Lys Gly Arg Leu Leu Lys Ala Gly Arg Gln Val Arg Glu Pro Gly Gln
500 505 510
Asp Leu Val Val Leu Pro Leu Ser Ile Thr Thr Asp Phe Ile Pro Ser
515 520 525
Phe Arg Leu Val Ala Tyr Tyr Thr Leu Ile Gly Ala Ser Gly Gln Arg
530 535 540
Glu Val Val Ala Asp Ser Val Trp Val Asp Val Lys Asp Ser Cys Val
545 550 555 560
Gly Ser Leu Val Val Lys Ser Gly Gln Ser Glu Asp Arg Gln Pro Val
565 570 575
Pro Gly Gln Gln Met Thr Leu Lys Ile Glu Gly Asp His Gly Ala Arg
580 585 590
Val Val Leu Val Ala Val Asp Lys Gly Val Phe Val Leu Asn Lys Lys
595 600 605
Asn Lys Leu Thr Gln Ser Lys Ile Trp Asp Val Val Glu Lys Ala Asp
610 615 620
Ile Gly Cys Thr Pro Gly Ser Gly Lys Asp Tyr Ala Gly Val Phe Ser
625 630 635 640
Asp Ala Gly Leu Thr Phe Thr Ser Ser Ser Gly Gln Gln Thr Ala Gln
645 650 655
Arg Ala Glu Leu Gln Cys Pro Gln Pro Ala Ala Arg Arg Arg Arg Ser
660 665 670
Val Gln Leu Thr Glu Lys Arg Met Asp Lys Ile Ser Thr Lys Leu Met
675 680 685
Asn Ile Phe Leu Lys Asp Ser Ile Thr Thr Trp Glu Ile Leu Ala Val
690 695 700
Ser Met Ser Asp Lys Lys Gly Ile Cys Val Ala Asp Pro Phe Glu Val
705 710 715 720

```
Thr Val Met Gln Asp Phe Phe Ile Asp Leu Arg Leu Pro Tyr Ser Val
                725                 730                 735
Val Arg Asn Glu Gln Val Glu Ile Arg Ala Val Leu Tyr Asn Tyr Arg
            740                 745                 750
Gln Asn Gln Glu Leu Lys Val Arg Val Glu Leu Leu His Asn Pro Ala
        755                 760                 765
Phe Cys Ser Leu Ala Thr Thr Lys Arg Arg His Gln Gln Thr Val Thr
    770                 775                 780
Ile Pro Pro Lys Ser Ser Leu Ser Val Pro Tyr Val Ile Val Pro Leu
785                 790                 795                 800
Lys Thr Gly Leu Val Gly Lys Tyr Pro Lys Glu Leu Arg Lys Cys Cys
                805                 810                 815
Glu Asp Gly Met Arg Glu Asn Pro Met Arg Phe Ser Cys Gln Arg Arg
            820                 825                 830
Thr Arg Phe Ile Ser Leu Gly Glu Ala Cys Lys Lys Val Phe Leu Asp
        835                 840                 845
Cys Cys Asn Tyr Ile Thr Glu Leu Arg Arg Gln His Ala Arg Ala Ser
    850                 855                 860
His Leu Gly Leu Ala Arg Ser Asn Leu Asp Glu Asp Ile Ile Ala Glu
865                 870                 875                 880
Glu Asn Ile Val Ser Arg Ser Glu Phe Pro Glu Ser Trp Leu Trp Asn
                885                 890                 895
Val Glu Asp Leu Lys Glu Pro Pro Lys Asn Gly Gln Glu Val Glu Val
            900                 905                 910
Lys Ala Ala Val Tyr His His Phe Ile Ser Asp Gly Val Arg Lys Ser
        915                 920                 925
Leu Lys Val Val Pro Glu Gly Ile Arg Met Asn Lys Thr Val Ala Val
    930                 935                 940
Arg Thr Leu Asp Pro Glu Arg Leu Gly Arg Glu Gly Val Gln Lys Glu
945                 950                 955                 960
Asp Ile Pro Pro Ala Asp Leu Ser Asp Gln Val Pro Asp Thr Glu Ser
                965                 970                 975
Glu Thr Arg Ile Leu Leu Gln Gly Thr Pro Val Ala Gln Met Thr Glu
            980                 985                 990
Asp Ala Val Asp Ala Glu Arg Leu Lys His Leu Ile Val Thr Pro Ser
        995                1000                1005
Gly Cys Gly Glu Gln Asn Met Ile Gly Met Thr Pro Thr Val Ile Ala
   1010                1015                1020
Val His Tyr Leu Asp Glu Thr Glu Gln Trp Glu Lys Phe Gly Leu Glu
025                1030                1035                1040
Lys Arg Gln Gly Ala Leu Glu Leu Ile Lys Lys Gly Tyr Thr Gln Gln
               1045                1050                1055
Leu Ala Phe Arg Gln Pro Ser Ser Ala Phe Ala Ala Phe Val Lys Arg
           1060                1065                1070
Ala Pro Ser Thr Trp Leu Thr Ala Tyr Val Val Lys Val Phe Ser Leu
       1075                1080                1085
Ala Val Asn Leu Ile Ala Ile Asp Ser Gln Val Leu Cys Gly Ala Val
   1090                1095                1100
Lys Trp Leu Ile Leu Glu Lys Gln Lys Pro Asp Gly Val Phe Gln Glu
105                1110                1115                1120
Asp Ala Pro Val Ile His Gln Glu Met Ile Gly Gly Leu Arg Asn Asn
               1125                1130                1135
Asn Glu Lys Asp Met Ala Leu Thr Ala Phe Val Leu Ile Ser Leu Gln
           1140                1145                1150
Glu Ala Lys Asp Ile Cys Glu Glu Gln Val Asn Ser Leu Pro Gly Ser
       1155                1160                1165
Ile Thr Lys Ala Gly Asp Phe Leu Glu Ala Asn Tyr Met Asn Leu Gln
   1170                1175                1180
Arg Ser Tyr Thr Val Ala Ile Ala Gly Tyr Ala Leu Ala Gln Met Gly
185                1190                1195                1200
Arg Leu Lys Gly Pro Leu Leu Asn Lys Phe Leu Thr Thr Ala Lys Asp
               1205                1210                1215
Lys Asn Arg Trp Glu Asp Pro Gly Lys Gln Leu Tyr Asn Val Glu Ala
           1220                1225                1230
Thr Ser Tyr Ala Leu Leu Ala Leu Leu Gln Leu Lys Asp Phe Asp Phe
       1235                1240                1245
Val Pro Pro Val Val Arg Trp Leu Asn Glu Gln Arg Tyr Tyr Gly Gly
   1250                1255                1260
Gly Tyr Gly Ser Thr Gln Ala Thr Phe Met Val Phe Gln Ala Leu Ala
265                1270                1275                1280
```

```
Gln Tyr Gln Lys Asp Ala Pro Asp His Gln Glu Leu Asn Leu Asp Val
                1285                1290                1295
Ser Leu Gln Leu Pro Ser Arg Ser Ser Lys Ile Thr His Arg Ile His
            1300                1305                1310
Trp Glu Ser Ala Ser Leu Leu Arg Ser Glu Glu Thr Lys Glu Asn Glu
        1315                1320                1325
Gly Phe Thr Val Thr Ala Glu Gly Lys Gly Gln Gly Thr Leu Ser Val
    1330                1335                1340
Val Thr Met Tyr His Ala Lys Ala Lys Asp Gln Leu Thr Cys Asn Lys
345                 1350                1355                1360
Phe Asp Leu Lys Val Thr Ile Lys Pro Ala Pro Glu Thr Glu Lys Arg
                1365                1370                1375
Pro Gln Asp Ala Lys Asn Thr Met Ile Leu Glu Ile Cys Thr Arg Tyr
            1380                1385                1390
Arg Gly Asp Gln Asp Ala Thr Met Ser Ile Leu Asp Ile Ser Met Met
        1395                1400                1405
Thr Gly Phe Ala Pro Asp Thr Asp Asp Leu Lys Gln Leu Ala Asn Gly
    1410                1415                1420
Val Asp Arg Tyr Ile Ser Lys Tyr Glu Leu Asp Lys Ala Phe Ser Asp
425                 1430                1435                1440
Arg Asn Thr Leu Ile Ile Tyr Leu Asp Lys Val Ser His Ser Glu Asp
                1445                1450                1455
Asp Cys Leu Ala Phe Lys Val His Gln Tyr Phe Asn Val Glu Leu Ile
            1460                1465                1470
Gln Pro Gly Ala Val Lys Val Tyr Ala Tyr Tyr Asn Leu Glu Glu Ser
        1475                1480                1485
Cys Thr Arg Phe Tyr His Pro Glu Lys Glu Asp Gly Lys Leu Asn Lys
    1490                1495                1500
Leu Cys Arg Asp Glu Leu Cys Arg Cys Ala Glu Glu Asn Cys Phe Ile
505                 1510                1515                1520
Gln Lys Ser Asp Asp Lys Val Thr Leu Glu Glu Arg Leu Asp Lys Ala
                1525                1530                1535
Cys Glu Pro Gly Val Asp Tyr Val Tyr Lys Thr Arg Leu Val Lys Val
            1540                1545                1550
Gln Leu Ser Asn Asp Phe Asp Glu Tyr Ile Met Ala Ile Glu Gln Thr
        1555                1560                1565
Ile Lys Ser Gly Ser Asp Glu Val Gln Val Gly Gln Gln Arg Thr Phe
    1570                1575                1580
Ile Ser Pro Ile Lys Cys Arg Glu Ala Leu Lys Leu Glu Glu Lys Lys
585                 1590                1595                1600
His Tyr Leu Met Trp Gly Leu Ser Ser Asp Phe Trp Gly Glu Lys Pro
                1605                1610                1615
Asn Leu Ser Tyr Ile Ile Gly Lys Asp Thr Trp Val Glu His Trp Pro
            1620                1625                1630
Glu Glu Asp Glu Cys Gln Asp Glu Glu Asn Gln Lys Gln Cys Gln Asp
        1635                1640                1645
Leu Gly Ala Phe Thr Glu Ser Met Val Val Phe Gly Cys Pro Asn
    1650                1655                1660
```

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 1663 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Gly Pro Ala Ser Gly Ser Gln Leu Leu Val Leu Leu Leu Leu Leu
 1               5                  10                  15
Ala Ser Ser Pro Leu Ala Leu Gly Ile Pro Met Tyr Ser Ile Ile Thr
            20                  25                  30
Pr  Asn Val Leu Arg Leu Glu Ser Glu Glu Thr Ile Val Leu Glu Ala
```

```
        35                  40                  45
His Asp Ala Gln Gly Asp Ile Pro Val Thr Val Thr Val Gln Asp Phe
    50                  55                  60
Leu Lys Arg Gln Val Leu Thr Ser Glu Lys Thr Val Leu Thr Gly Ala
65                  70                  75                  80
Ser Gly His Leu Arg Ser Val Ser Ile Lys Ile Pro Ala Ser Lys Glu
                85                  90                  95
Phe Asn Ser Asp Lys Glu Gly His Lys Tyr Val Thr Val Val Ala Asn
            100                 105                 110
Phe Gly Glu Thr Val Val Glu Lys Ala Val Met Val Ser Phe Gln Ser
        115                 120                 125
Gly Tyr Leu Phe Ile Gln Thr Asp Lys Thr Ile Tyr Thr Pro Gly Ser
    130                 135                 140
Thr Val Leu Tyr Arg Ile Phe Thr Val Asp Asn Asn Leu Leu Pro Val
145                 150                 155                 160
Gly Lys Thr Val Val Ile Leu Ile Glu Thr Pro Asp Gly Ile Pro Val
                165                 170                 175
Lys Arg Asp Ile Leu Ser Ser Asn Asn Gln His Gly Ile Leu Pro Leu
            180                 185                 190
Ser Trp Asn Ile Pro Glu Leu Val Asn Met Gly Gln Trp Lys Ile Arg
        195                 200                 205
Ala Phe Tyr Glu His Ala Pro Lys Gln Ile Phe Ser Ala Glu Phe Glu
    210                 215                 220
Val Lys Glu Tyr Val Leu Pro Ser Phe Glu Val Arg Val Glu Pro Thr
225                 230                 235                 240
Glu Thr Phe Tyr Tyr Ile Asp Asp Pro Asn Gly Leu Glu Val Ser Ile
                245                 250                 255
Ile Ala Lys Phe Leu Tyr Gly Lys Asn Val Asp Gly Thr Ala Phe Val
            260                 265                 270
Ile Phe Gly Val Gln Asp Gly Asp Lys Lys Ile Ser Leu Ala His Ser
        275                 280                 285
Leu Thr Arg Val Val Ile Glu Asp Gly Val Gly Asp Ala Val Leu Thr
    290                 295                 300
Arg Lys Val Leu Met Glu Gly Val Arg Pro Ser Asn Ala Asp Ala Leu
305                 310                 315                 320
Val Gly Lys Ser Leu Tyr Val Ser Val Thr Val Ile Leu His Ser Gly
                325                 330                 335
Ser Asp Met Val Glu Ala Glu Arg Ser Gly Ile Pro Ile Val Thr Ser
            340                 345                 350
Pro Tyr Gln Ile His Phe Thr Lys Thr Pro Lys Phe Phe Lys Pro Ala
        355                 360                 365
Met Pro Phe Asp Leu Met Val Phe Val Thr Asn Pro Asp Gly Ser Pro
    370                 375                 380
Ala Ser Lys Val Leu Val Val Thr Gln Gly Ser Asn Ala Lys Ala Leu
385                 390                 395                 400
Thr Gln Asp Asp Gly Val Ala Lys Leu Ser Ile Asn Thr Pro Asn Ser
                405                 410                 415
Arg Gln Pro Leu Thr Ile Thr Val Arg Thr Lys Lys Asp Thr Leu Pro
            420                 425                 430
Glu Ser Arg Gln Ala Thr Lys Thr Met Glu Ala His Pro Tyr Ser Thr
        435                 440                 445
Met His Asn Ser Asn Asn Tyr Leu His Leu Ser Val Ser Arg Met Glu
    450                 455                 460
Leu Lys Pro Gly Asp Asn Leu Asn Val Asn Phe His Leu Arg Thr Asp
465                 470                 475                 480
Pro Gly His Glu Ala Lys Ile Arg Tyr Tyr Thr Tyr Leu Val Met Asn
                485                 490                 495
Lys Gly Lys Leu Leu Lys Ala Gly Arg Gln Val Arg Glu Pro Gly Gln
            500                 505                 510
Asp Leu Val Val Leu Ser Leu Pro Ile Thr Pro Glu Phe Ile Pro Ser
        515                 520                 525
Phe Arg Leu Val Ala Tyr Tyr Thr Leu Ile Gly Ala Ser Gly Gln Arg
    530                 535                 540
Glu Val Val Ala Asp Ser Val Trp Val Asp Val Lys Asp Ser Cys Ile
545                 550                 555                 560
Gly Thr Leu Val Val Lys Gly Asp Pro Arg Asp Asn His Leu Ala Pro
                565                 570                 575
Gly Gln Gln Thr Thr Leu Arg Ile Glu Gly Asn Gln Gly Ala Arg Val
            580                 585                 590
Gly Leu Val Ala Val Asp Lys Gly Val Phe Val Leu Asn Lys Lys Asn
```

595 600 605
Lys Leu Thr Gln Ser Lys Ile Trp Asp Val Val Glu Lys Ala Asp Ile
610 615 620
Gly Cys Thr Pro Gly Ser Gly Lys Asn Tyr Ala Gly Val Phe Met Asp
625 630 635 640
Ala Gly Leu Ala Phe Lys Thr Ser Gln Gly Leu Gln Thr Glu Gln Arg
645 650 655
Ala Asp Leu Glu Cys Thr Lys Pro Ala Ala Arg Arg Arg Arg Ser Val
660 665 670
Gln Leu Met Glu Arg Arg Met Asp Lys Ala Gly Gln Tyr Thr Asp Lys
675 680 685
Gly Leu Arg Lys Cys Cys Glu Asp Gly Met Arg Asp Ile Pro Met Arg
690 695 700
Tyr Ser Cys Gln Arg Arg Ala Arg Leu Ile Thr Gln Gly Glu Asn Cys
705 710 715 720
Ile Lys Ala Phe Ile Asp Cys Cys Asn His Ile Thr Lys Leu Arg Glu
725 730 735
Gln His Arg Arg Asp His Val Leu Gly Leu Ala Arg Ser Glu Leu Glu
740 745 750
Glu Asp Ile Ile Pro Glu Glu Asp Ile Ile Ser Arg Ser His Phe Pro
755 760 765
Gln Ser Trp Leu Trp Thr Ile Glu Glu Leu Lys Glu Pro Glu Lys Asn
770 775 780
Gly Ile Ser Thr Lys Val Met Asn Ile Phe Leu Lys Asp Ser Ile Thr
785 790 795 800
Thr Trp Glu Ile Leu Ala Val Ser Leu Ser Asp Lys Lys Gly Ile Cys
805 810 815
Val Ala Asp Pro Tyr Glu Ile Arg Val Met Gln Asp Phe Phe Ile Asp
820 825 830
Leu Arg Leu Pro Tyr Ser Val Val Arg Asn Glu Gln Val Glu Ile Arg
835 840 845
Ala Val Leu Phe Asn Tyr Arg Glu Gln Gln Glu Leu Lys Val Arg Val
850 855 860
Glu Leu Leu His Asn Pro Ala Phe Cys Ser Met Ala Thr Ala Lys Asn
865 870 875 880
Arg Tyr Phe Gln Thr Ile Lys Ile Pro Pro Lys Ser Ser Val Ala Val
885 890 895
Pro Tyr Val Ile Val Pro Leu Lys Ile Gly Gln Gln Glu Val Glu Val
900 905 910
Lys Ala Ala Val Phe Asn His Phe Ile Ser Asp Gly Val Lys Lys Thr
915 920 925
Leu Lys Val Val Pro Glu Gly Met Arg Ile Asn Lys Thr Val Ala Ile
930 935 940
His Thr Leu Asp Pro Glu Lys Leu Gly Gln Gly Gly Val Gln Lys Val
945 950 955 960
Asp Val Pro Ala Ala Asp Leu Ser Asp Gln Val Pro Asp Thr Asp Ser
965 970 975
Glu Thr Arg Ile Ile Leu Gln Gly Ser Pro Val Val Gln Met Ala Glu
980 985 990
Asp Ala Val Asp Gly Glu Arg Leu Lys His Leu Ile Val Thr Pro Ala
995 1000 1005
Gly Cys Gly Glu Gln Asn Met Ile Gly Met Thr Pro Thr Val Ile Ala
1010 1015 1020
Val His Tyr Leu Asp Gln Thr Glu Gln Trp Glu Lys Phe Gly Ile Glu
025 1030 1035 1040
Lys Arg Gln Glu Ala Leu Glu Leu Ile Lys Lys Gly Tyr Thr Gln Gln
1045 1050 1055
Leu Ala Phe Lys Gln Pro Ser Ser Ala Tyr Ala Ala Phe Asn Asn Arg
1060 1065 1070
Pro Pro Ser Thr Trp Leu Thr Ala Tyr Val Val Lys Val Phe Ser Leu
1075 1080 1085
Ala Ala Asn Leu Ile Ala Ile Asp Ser His Val Leu Cys Gly Ala Val
1090 1095 1100
Lys Trp Leu Ile Leu Glu Lys Gln Lys Pro Asp Gly Val Phe Gln Glu
105 1110 1115 1120
Asp Gly Pro Val Ile His Gln Glu Met Ile Gly Gly Phe Arg Asn Ala
1125 1130 1135
Lys Glu Ala Asp Val Ser Leu Thr Ala Phe Val Leu Ile Ala Leu Gln
1140 1145 1150
Glu Ala Arg Asp Ile Cys Glu Gly Gln Val Asn Ser Leu Pro Gly Ser

```
        1155                1160                1165
Ile Asn Lys Ala Gly Glu Tyr Ile Glu Ala Ser Tyr Met Asn Leu Gln
    1170                1175                1180
Arg Pro Tyr Thr Val Ala Ile Ala Gly Tyr Ala Leu Ala Leu Met Asn
185                 1190                1195                1200
Lys Leu Glu Glu Pro Tyr Leu Gly Lys Phe Leu Asn Thr Ala Lys Asp
                1205                1210                1215
Arg Asn Arg Trp Glu Glu Pro Asp Gln Gln Leu Tyr Asn Val Glu Ala
            1220                1225                1230
Thr Ser Tyr Ala Leu Leu Ala Leu Leu Leu Leu Lys Asp Phe Asp Ser
        1235                1240                1245
Val Pro Pro Val Val Arg Trp Leu Asn Glu Gln Arg Tyr Tyr Gly Gly
    1250                1255                1260
Gly Tyr Gly Ser Thr Gln Ala Thr Phe Met Val Phe Gln Ala Leu Ala
265                 1270                1275                1280
Gln Tyr Gln Thr Asp Val Pro Asp His Lys Asp Leu Asn Met Asp Val
                1285                1290                1295
Ser Phe His Leu Pro Ser Arg Ser Ser Ala Thr Thr Phe Arg Leu Leu
            1300                1305                1310
Trp Glu Asn Gly Asn Leu Leu Arg Ser Glu Glu Thr Lys Gln Asn Glu
        1315                1320                1325
Ala Phe Ser Leu Thr Ala Lys Gly Lys Gly Arg Gly Thr Leu Ser Val
    1330                1335                1340
Val Ala Val Tyr His Ala Lys Leu Lys Ser Lys Val Thr Cys Lys Lys
345                 1350                1355                1360
Phe Asp Leu Arg Val Ser Ile Arg Pro Ala Pro Glu Thr Ala Lys Lys
                1365                1370                1375
Pro Glu Glu Ala Lys Asn Thr Met Phe Leu Glu Ile Cys Thr Lys Tyr
            1380                1385                1390
Leu Gly Asp Val Asp Ala Thr Met Ser Ile Leu Asp Ile Ser Met Met
        1395                1400                1405
Thr Gly Phe Ala Pro Asp Thr Lys Asp Leu Glu Leu Leu Ala Ser Gly
    1410                1415                1420
Val Asp Arg Tyr Ile Ser Lys Tyr Glu Met Asn Lys Ala Phe Ser Asn
425                 1430                1435                1440
Lys Asn Thr Leu Ile Ile Tyr Leu Glu Lys Ile Ser His Thr Glu Glu
                1445                1450                1455
Asp Cys Leu Thr Phe Lys Val His Gln Tyr Phe Asn Val Gly Leu Ile
            1460                1465                1470
Gln Pro Gly Ser Val Lys Val Tyr Ser Tyr Tyr Asn Leu Glu Glu Ser
        1475                1480                1485
Cys Thr Arg Phe Tyr His Pro Glu Lys Asp Asp Gly Met Leu Ser Lys
    1490                1495                1500
Leu Cys His Ser Glu Met Cys Arg Cys Ala Glu Glu Asn Cys Phe Met
505                 1510                1515                1520
Gln Gln Ser Gln Glu Lys Ile Asn Leu Asn Val Arg Leu Asp Lys Ala
                1525                1530                1535
Cys Glu Pro Gly Val Asp Tyr Val Tyr Lys Thr Glu Leu Thr Asn Ile
            1540                1545                1550
Lys Leu Leu Asp Asp Phe Asp Glu Tyr Thr Met Thr Ile Gln Gln Val
        1555                1560                1565
Ile Lys Ser Gly Ser Asp Glu Val Gln Ala Gly Gln Gln Arg Lys Phe
    1570                1575                1580
Ile Ser His Ile Lys Cys Arg Asn Ala Leu Lys Leu Gln Lys Gly Lys
585                 1590                1595                1600
Lys Tyr Leu Met Trp Gly Leu Ser Ser Asp Leu Trp Gly Glu Lys Pro
                1605                1610                1615
Asn Thr Ser Tyr Ile Ile Gly Lys Asp Thr Trp Val Glu His Trp Pro
            1620                1625                1630
Glu Ala Glu Glu Cys Gln Asp Gln Lys Tyr Gln Lys Gln Cys Glu Glu
        1635                1640                1645
Leu Gly Ala Phe Thr Glu Ser Met Val Val Tyr Gly Cys Pro Asn
    1650                1655                1660
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1744 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
Met Arg Leu Leu Trp Gly Leu Ile Trp Ala Ser Ser Phe Phe Thr Leu
 1               5                  10                  15
Ser Leu Gln Lys Pro Arg Leu Leu Leu Phe Ser Pro Ser Val Val His
            20                  25                  30
Leu Gly Val Pro Leu Ser Val Gly Val Gln Leu Gln Asp Val Pro Arg
        35                  40                  45
Gly Gln Val Val Lys Gly Ser Val Phe Leu Arg Asn Pro Ser Arg Asn
    50                  55                  60
Asn Val Pro Cys Ser Pro Lys Val Asp Phe Thr Leu Ser Ser Glu Arg
65                  70                  75                  80
Asp Phe Ala Leu Leu Ser Leu Gln Val Pro Leu Lys Asp Ala Lys Ser
                85                  90                  95
Cys Gly Leu His Gln Leu Leu Arg Gly Pro Glu Val Gln Leu Val Ala
            100                 105                 110
His Ser Pro Trp Leu Lys Asp Ser Leu Ser Arg Thr Thr Asn Ile Gln
        115                 120                 125
Gly Ile Asn Leu Leu Phe Ser Ser Arg Arg Gly His Leu Phe Leu Gln
    130                 135                 140
Thr Asp Gln Pro Ile Tyr Asn Pro Gly Gln Arg Val Arg Tyr Arg Val
145                 150                 155                 160
Phe Ala Leu Asp Gln Lys Met Arg Pro Ser Thr Asp Thr Ile Thr Val
                165                 170                 175
Met Val Glu Asn Ser His Gly Leu Arg Val Arg Lys Lys Glu Val Tyr
            180                 185                 190
Met Pro Ser Ser Ile Phe Gln Asp Asp Phe Val Ile Pro Asp Ile Ser
        195                 200                 205
Glu Pro Gly Thr Trp Lys Ile Ser Ala Arg Phe Ser Asp Gly Leu Glu
    210                 215                 220
Ser Asn Ser Ser Thr Gln Phe Glu Val Lys Lys Tyr Val Leu Pro Asn
225                 230                 235                 240
Phe Glu Val Lys Ile Thr Pro Gly Lys Pro Tyr Ile Leu Thr Val Pro
                245                 250                 255
Gly His Leu Asp Glu Met Gln Leu Asp Ile Gln Ala Arg Tyr Ile Tyr
            260                 265                 270
Gly Lys Pro Val Gln Gly Val Ala Tyr Val Arg Phe Gly Leu Leu Asp
        275                 280                 285
Glu Asp Gly Lys Lys Thr Phe Phe Arg Gly Leu Glu Ser Gln Thr Lys
    290                 295                 300
Leu Val Asn Gly Gln Ser His Ile Ser Leu Ser Lys Ala Glu Phe Gln
305                 310                 315                 320
Asp Ala Leu Glu Lys Leu Asn Met Gly Ile Thr Asp Leu Gln Gly Leu
                325                 330                 335
Arg Leu Tyr Val Ala Ala Ala Ile Ile Glu Ser Pro Gly Gly Glu Met
            340                 345                 350
Glu Glu Ala Glu Leu Thr Ser Trp Tyr Phe Val Ser Ser Pro Phe Ser
        355                 360                 365
Leu Asp Leu Ser Lys Thr Lys Arg His Leu Val Pro Gly Ala Pro Phe
    370                 375                 380
Leu Leu Gln Ala Leu Val Arg Glu Met Ser Gly Ser Pro Ala Ser Gly
385                 390                 395                 400
Ile Pro Val Lys Val Ser Ala Thr Val Ser Ser Pro Gly Ser Val Pro
                405                 410                 415
Glu Ala Gln Asp Ile Gln Gln Asn Thr Asp Gly Ser Gly Gln Val Ser
            420                 425                 430
Ile Pro Ile Ile Ile Pro Gln Thr Ile Ser Glu Leu Gln Leu Ser Val
        435                 440                 445
Ser Ala Gly Ser Pro His Pro Ala Ile Ala Arg Leu Thr Val Ala Ala
    450                 455                 460
Pro Pro Ser Gly Gly Pro Gly Phe Leu Ser Ile Glu Arg Pro Asp Ser
465                 470                 475                 480
```

Arg Pro Pro Arg Val Gly Asp Thr Leu Asn Leu Asn Leu Arg Ala Val
485 490 495
Gly Ser Gly Ala Thr Phe Ser His Tyr Tyr Tyr Met Ile Leu Ser Arg
500 505 510
Gly Gln Ile Val Phe Met Asn Arg Glu Pro Lys Arg Thr Leu Thr Ser
515 520 525
Val Ser Val Phe Val Asp His His Leu Ala Pro Ser Phe Tyr Phe Val
530 535 540
Ala Phe Tyr Tyr His Gly Asp His Pro Val Ala Asn Ser Leu Arg Val
545 550 555 560
Asp Val Gln Ala Gly Ala Cys Glu Gly Lys Leu Glu Leu Ser Val Asp
565 570 575
Gly Ala Lys Gln Tyr Arg Asn Gly Glu Ser Val Lys Leu His Leu Glu
580 585 590
Thr Asp Ser Leu Ala Leu Val Ala Leu Gly Ala Leu Asp Thr Ala Leu
595 600 605
Tyr Ala Ala Gly Ser Lys Ser His Lys Pro Leu Asn Met Gly Lys Val
610 615 620
Phe Glu Ala Met Asn Ser Tyr Asp Leu Gly Cys Gly Pro Gly Gly Gly
625 630 635 640
Asp Ser Ala Leu Gln Val Phe Gln Ala Ala Gly Leu Ala Phe Ser Asp
645 650 655
Gly Asp Gln Trp Thr Leu Ser Arg Lys Arg Leu Ser Cys Pro Lys Glu
660 665 670
Lys Thr Thr Arg Lys Lys Arg Asn Val Asn Phe Gln Lys Ala Ile Asn
675 680 685
Glu Lys Leu Gly Gln Tyr Ala Ser Pro Thr Ala Lys Arg Cys Cys Gln
690 695 700
Asp Gly Val Thr Arg Leu Pro Met Met Arg Ser Cys Glu Gln Arg Ala
705 710 715 720
Ala Arg Val Gln Gln Pro Asp Cys Arg Glu Pro Phe Leu Ser Cys Cys
725 730 735
Gln Phe Ala Glu Ser Leu Arg Lys Lys Ser Arg Asp Lys Gly Gln Ala
740 745 750
Gly Leu Gln Arg Ala Leu Glu Ile Leu Gln Glu Glu Asp Leu Ile Asp
755 760 765
Glu Asp Asp Ile Pro Val Arg Ser Phe Phe Pro Glu Asn Trp Leu Trp
770 775 780
Arg Val Glu Thr Val Asp Arg Phe Gln Ile Leu Thr Leu Trp Leu Pro
785 790 795 800
Asp Ser Leu Thr Thr Trp Glu Ile His Gly Leu Ser Leu Ser Lys Thr
805 810 815
Lys Gly Leu Cys Val Ala Thr Pro Val Gln Leu Arg Val Phe Arg Glu
820 825 830
Phe His Leu His Leu Arg Leu Pro Met Ser Val Arg Arg Phe Glu Gln
835 840 845
Leu Glu Leu Arg Pro Val Leu Tyr Asn Tyr Leu Asp Lys Asn Leu Thr
850 855 860
Val Ser Val His Val Ser Pro Val Glu Gly Leu Cys Leu Ala Gly Gly
865 870 875 880
Gly Gly Leu Ala Gln Gln Val Leu Val Pro Ala Gly Ser Ala Arg Pro
885 890 895
Val Ala Phe Ser Val Val Pro Thr Ala Ala Ala Ala Val Ser Leu Lys
900 905 910
Val Val Ala Arg Gly Ser Phe Glu Phe Pro Val Gly Asp Ala Val Ser
915 920 925
Lys Val Leu Gln Ile Glu Lys Glu Gly Ala Ile His Arg Glu Glu Leu
930 935 940
Val Tyr Glu Leu Asn Pro Leu Asp His Arg Gly Arg Thr Leu Glu Ile
945 950 955 960
Pro Gly Asn Ser Asp Pro Asn Met Ile Pro Asp Gly Asp Phe Asn Ser
965 970 975
Tyr Val Arg Val Thr Ala Ser Asp Pro Leu Asp Thr Leu Gly Ser Glu
980 985 990
Gly Ala Leu Ser Pro Gly Gly Val Ala Ser L . Leu Arg Leu Pro Arg
995 1000 1005
Gly Cys Gly Glu Gln Thr Met Ile Tyr Leu Ala Pro Thr Leu Ala Ala
1010 1015 1020
Ser Arg Tyr Leu Asp Lys Thr Glu Gln Trp Ser Thr Leu Pro Pro Glu
025 1030 1035 1040

Thr Lys Asp His Ala Val Asp Leu Ile Gln Lys Gly Tyr Met Arg Ile
1045 1050 1055
Gln Gln Phe Arg Lys Ala Asp Gly Ser Tyr Ala Ala Trp Leu Ser Arg
1060 1065 1070
Asp Ser Ser Thr Trp Leu Thr Ala Phe Val Leu Lys Val Leu Ser Leu
1075 1080 1085
Ala Gln Glu Gln Val Gly Gly Ser Pro Glu Lys Leu Gln Glu Thr Ser
1090 1095 1100
Asn Trp Leu Leu Ser Gln Gln Gln Ala Asp Gly Ser Phe Gln Asp Pro
105 1110 1115 1120
Cys Pro Val Leu Asp Arg Ser Met Gln Gly Gly Leu Val Gly Asn Asp
1125 1130 1135
Glu Thr Val Ala Leu Thr Ala Phe Val Thr Ile Ala Leu His His Gly
1140 1145 1150
Leu Ala Val Phe Gln Asp Glu Gly Ala Glu Pro Leu Lys Gln Arg Val
1155 1160 1165
Glu Ala Ser Ile Ser Lys Ala Asn Ser Phe Leu Gly Glu Lys Ala Ser
1170 1175 1180
Ala Gly Leu Leu Gly Ala His Ala Ala Ala Ile Thr Ala Tyr Ala Leu
185 1190 1195 1200
Ser Leu Thr Lys Ala Pro Val Asp Leu Leu Gly Val Ala His Asn Asn
1205 1210 1215
Leu Met Ala Met Ala Gln Glu Thr Gly Asp Asn Leu Tyr Trp Gly Ser
1220 1225 1230
Val Thr Gly Ser Gln Ser Asn Ala Val Ser Pro Thr Pro Ala Pro Arg
1235 1240 1245
Asn Pro Ser Asp Pro Met Pro Gln Ala Pro Ala Leu Trp Ile Glu Thr
1250 1255 1260
Thr Ala Tyr Ala Leu Leu His Leu Leu Leu His Glu Gly Lys Ala Glu
265 1270 1275 1280
Met Ala Asp Gln Ala Ser Ala Trp Leu Thr Arg Gln Gly Ser Phe Gln
1285 1290 1295
Gly Gly Phe Arg Ser Thr Gln Asp Thr Val Ile Ala Leu Asp Ala Leu
1300 1305 1310
Ser Ala Tyr Trp Ile Ala Ser His Thr Thr Glu Glu Arg Gly Leu Asn
1315 1320 1325
Val Thr Leu Ser Ser Thr Gly Arg Asn Gly Phe Lys Ser His Ala Leu
1330 1335 1340
Gln Leu Asn Asn Arg Gln Ile Arg Gly Leu Glu Glu Glu Leu Gln Phe
345 1350 1355 1360
Ser Leu Gly Ser Lys Ile Asn Val Lys Val Gly Gly Asn Ser Lys Gly
1365 1370 1375
Thr Leu Lys Val Leu Arg Thr Tyr Asn Val Leu Asp Met Lys Asn Thr
1380 1385 1390
Thr Cys Gln Asp Leu Gln Ile Glu Val Thr Val Lys Gly His Val Glu
1395 1400 1405
Tyr Thr Met Glu Ala Asn Glu Asp Tyr Glu Asp Tyr Glu Tyr Asp Glu
1410 1415 1420
Leu Pro Ala Lys Asp Asp Pro Asp Ala Pro Leu Gln Pro Val Thr Pro
425 1430 1435 1440
Leu Gln Leu Phe Glu Gly Arg Arg Asn Arg Arg Arg Arg Glu Ala Pro
1445 1450 1455
Lys Val Val Glu Glu Gln Glu Ser Arg Val His Tyr Thr Val Cys Ile
1460 1465 1470
Trp Arg Asn Gly Lys Val Gly Leu Ser Gly Met Ala Ile Ala Asp Val
1475 1480 1485
Thr Leu Leu Ser Gly Phe His Ala Leu Arg Ala Asp Leu Glu Lys Leu
1490 1495 1500
Thr Ser Leu Ser Asp Arg Tyr Val Ser His Phe Glu Thr Glu Gly Pro
505 1510 1515 1520
His Val Leu Leu Tyr Phe Asp Ser Val Pro Thr Ser Arg Glu Cys Val
1525 1530 1535
Gly Phe Glu Ala Val Gln Glu Val Pro Val Gly Leu Val Gln Pro Ala
1540 1545 1550
Ser Ala Thr Leu Tyr Asp Tyr Tyr Asn Pro Glu Arg Arg Cys Ser Val
1555 1560 1565
Phe Tyr Gly Ala Pro Ser Lys Ser Arg Leu Leu Ala Thr Leu Cys Ser
1570 1575 1580
Ala Glu Val Cys Gln Cys Ala Glu Gly Lys Cys Pro Arg Gln Arg Arg
585 1590 1595 1600

Ala Leu Glu Arg Gly Leu Gln Asp Glu Asp Gly Tyr Arg Met Lys Phe
1605 1610 1615
Ala Cys Tyr Tyr Pro Arg Val Glu Tyr Gly Phe Gln Val Lys Val Leu
1620 1625 1630
Arg Glu Asp Ser Arg Ala Ala Phe Arg Leu Phe Glu Thr Lys Ile Thr
1635 1640 1645
Gln Val Leu His Phe Thr Lys Asp Val Lys Ala Ala Ala Asn Gln Met
1650 1655 1660
Arg Asn Phe Leu Val Arg Ala Ser Cys Arg Leu Arg Leu Glu Pro Gly
665 1670 1675 1680
Lys Glu Tyr Leu Ile Met Gly Leu Asp Gly Ala Thr Tyr Asp Leu Glu
1685 1690 1695
Gly His Pro Gln Tyr Leu Leu Asp Ser Asn Ser Trp Ile Glu Glu Met
1700 1705 1710
Pro Ser Glu Arg Leu Cys Arg Ser Thr Arg Gln Arg Ala Ala Cys Ala
1715 1720 1725
Gln Leu Asn Asp Phe Leu Gln Glu Tyr Gly Thr Gln Gly Cys Gln Val
1730 1735 1740

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1738 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Arg Leu Leu Trp Gly Leu Ala Trp Val Phe Ser Phe Cys Ala Ser
1 5 10 15
Ser Leu Gln Lys Pro Arg Leu Leu Leu Phe Ser Pro Ser Val Val Asn
20 25 30
Leu Gly Thr Pro Leu Ser Val Gly Val Gln Leu Leu Asp Ala Pro Pro
35 40 45
Gly Gln Glu Val Lys Gly Ser Val Phe Leu Arg Asn Pro Lys Gly Gly
50 55 60
Ser Cys Ser Pro Lys Lys Asp Phe Lys Leu Ser Ser Gly Asp Asp Phe
65 70 75 80
Val Leu Leu Ser Leu Glu Val Pro Leu Glu Asp Val Arg Ser Cys Gly
85 90 95
Leu Phe Asp Leu Arg Arg Ala Pro His Ile Gln Leu Val Ala Gln Ser
100 105 110
Pro Trp Leu Arg Asn Thr Ala Phe Lys Ala Thr Glu Thr Gln Gly Val
115 120 125
Asn Leu Leu Phe Ser Ser Arg Arg Gly His Ile Phe Val Gln Thr Asp
130 135 140
Gln Pro Ile Tyr Asn Pro Gly Gln Arg Val Arg Tyr Arg Val Phe Ala
145 150 155 160
Leu Asp Gln Lys Met Arg Pro Ser Thr Asp Phe Leu Thr Ile Thr Val
165 170 175
Glu Asn Ser His Gly Leu Arg Val Leu Lys Lys Glu Ile Phe Thr Ser
180 185 190
Thr Ser Ile Phe Gln Asp Ala Phe Thr Ile Pro Asp Ile Ser Glu Pro
195 200 205
Gly Thr Trp Lys Ile Ser Ala Arg Phe Ser Asp Gly Leu Glu Ser Asn
210 215 220
Arg Ser Thr His Phe Glu Val Lys Lys Tyr Val Leu Pro Asn Phe Glu
225 230 235 240
Val Lys Ile Thr Pro Trp Lys Pro Tyr Ile Leu Met Val Pro Ser Asn
245 250 255
Ser Asp Glu Ile Gln Leu Asp Ile Gln Ala Arg Tyr Ile Tyr Gly Lys
260 265 270
Pro Val Gln Gly Val Ala Tyr Thr Arg Phe Ala Leu Met Asp Glu Gln

```
            275                     280                     285
Gly Lys Arg Thr Phe Leu Arg Gly Leu Glu Thr Gln Ala Lys Leu Val
    290                     295                     300
Glu Gly Arg Thr His Ile Ser Ile Ser Lys Asp Gln Phe Gln Ala Ala
305                     310                     315                     320
Leu Asp Lys Ile Asn Ile Gly Val Arg Asp Leu Glu Gly Leu Arg Leu
                325                     330                     335
Tyr Ala Ala Thr Ala Val Ile Glu Ser Pro Gly Gly Glu Met Glu Glu
            340                     345                     350
Ala Glu Leu Thr Ser Trp Arg Phe Val Ser Ser Ala Phe Ser Leu Asp
        355                     360                     365
Leu Ser Arg Thr Lys Arg His Leu Val Pro Gly Ala His Phe Leu Leu
    370                     375                     380
Gln Ala Leu Val Gln Glu Met Ser Gly Ser Glu Ala Ser Asn Val Pro
385                     390                     395                     400
Val Lys Val Ser Ala Thr Leu Val Ser Gly Ser Asp Ser Gln Val Leu
                405                     410                     415
Asp Ile Gln Gln Ser Thr Asn Gly Ile Gly Gln Val Ser Ile Ser Phe
            420                     425                     430
Pro Ile Pro Pro Thr Val Thr Glu Leu Arg Leu Leu Val Ser Ala Gly
        435                     440                     445
Ser Leu Tyr Pro Ala Ile Ala Arg Leu Thr Val Gln Ala Pro Pro Ser
    450                     455                     460
Arg Gly Thr Gly Phe Leu Ser Ile Glu Pro Leu Asp Pro Arg Ser Pro
465                     470                     475                     480
Ser Val Gly Asp Thr Phe Ile Leu Asn Leu Gln Pro Val Gly Ile Pro
                485                     490                     495
Ala Pro Thr Phe Ser His Tyr Tyr Tyr Met Ile Ile Ser Arg Gly Gln
            500                     505                     510
Ile Met Ala Met Gly Arg Glu Pro Arg Lys Thr Val Thr Ser Val Ser
        515                     520                     525
Val Leu Val Asp His Gln Leu Ala Pro Ser Phe Tyr Phe Val Ala Tyr
    530                     535                     540
Phe Tyr His Gln Gly His Pro Val Ala Asn Ser Leu Leu Ile Asn Ile
545                     550                     555                     560
Gln Ser Arg Asp Cys Glu Gly Lys Leu Gln Leu Lys Val Asp Gly Ala
                565                     570                     575
Lys Glu Tyr Arg Asn Ala Asp Met Met Lys Leu Arg Ile Gln Thr Asp
            580                     585                     590
Ser Lys Ala Leu Val Ala Leu Gly Ala Val Asp Thr Ala Leu Tyr Ala
        595                     600                     605
Val Gly Gly Arg Ser His Lys Pro Leu Asp Met Ser Lys Val Phe Glu
    610                     615                     620
Val Ile Asn Ser Tyr Asn Val Gly Cys Gly Pro Gly Gly Gly Asp Asp
625                     630                     635                     640
Ala Leu Gln Val Phe Gln Asp Ala Gly Leu Ala Phe Ser Asp Gly Asp
                645                     650                     655
Arg Leu Thr Gln Thr Arg Glu Asp Leu Ser Cys Pro Lys Glu Lys Lys
            660                     665                     670
Ser Arg Gln Lys Arg Asn Val Asn Phe Gln Lys Ala Val Ser Glu Lys
        675                     680                     685
Leu Gly Gln Tyr Ser Ser Pro Asp Ala Lys Arg Cys Cys Gln Asp Gly
    690                     695                     700
Met Thr Lys Leu Pro Met Lys Arg Thr Cys Glu Gln Arg Ala Ala Arg
705                     710                     715                     720
Val Pro Gln Gln Ala Cys Arg Glu Pro Phe Leu Ser Cys Cys Lys Phe
                725                     730                     735
Ala Glu Asp Leu Arg Arg Asn Gln Thr Arg Ser Gln Ala His Leu Ala
            740                     745                     750
Arg Asn Asn His Asn Met Leu Gln Glu Glu Asp Leu Ile Asp Glu Asp
        755                     760                     765
Asp Ile Leu Val Arg Thr Ser Phe Pro Glu Asn Trp Leu Trp Arg Val
    770                     775                     780
Glu Pro Val Asp Ser Ser Lys Leu Leu Thr Val Trp Leu Pro Asp Ser
785                     790                     795                     800
Met Thr Thr Trp Glu Ile His Gly Val Ser Leu Ser Lys Ser Lys Gly
                805                     810                     815
Leu Cys Val Ala Lys Pro Thr Arg Val Arg Val Phe Arg Lys Phe His
            820                     825                     830
Leu His Leu Arg Leu Pro Ile Ser Ile Arg Arg Phe Glu Gln Phe Glu
```

```
        835                     840                     845
Leu Arg Pro Val Leu Tyr Asn Tyr Leu Asn Asp Asp Val Ala Val Ser
    850                     855                     860
Val His Val Thr Pro Val Glu Gly Leu Cys Leu Ala Gly Gly Gly Met
865                     870                     875                     880
Met Ala Gln Gln Val Thr Val Pro Ala Gly Ser Ala Arg Pro Val Ala
                885                     890                     895
Phe Ser Val Val Pro Thr Ala Ala Ala Asn Val Pro Leu Lys Val Val
            900                     905                     910
Ala Arg Gly Val Phe Asp Leu Gly Asp Ala Val Ser Lys Ile Leu Gln
        915                     920                     925
Ile Glu Lys Glu Gly Ala Ile His Arg Glu Glu Leu Val Tyr Asn Leu
    930                     935                     940
Asp Pro Leu Asn Asn Leu Gly Arg Thr Leu Glu Ile Pro Gly Ser Ser
945                     950                     955                     960
Asp Pro Asn Ile Val Pro Asp Gly Asp Phe Ser Ser Leu Val Arg Val
                965                     970                     975
Thr Ala Ser Glu Pro Leu Glu Thr Met Gly Ser Glu Gly Ala Leu Ser
            980                     985                     990
Pro Gly Gly Val Ala Ser Leu Leu Arg Leu Pro Gln Gly Cys Ala Glu
        995                    1000                    1005
Gln Thr Met Ile Tyr Leu Ala Pro Thr Leu Thr Ala Ser Asn Tyr Leu
   1010                    1015                    1020
Asp Arg Thr Glu Gln Trp Ser Lys Leu Ser Pro Glu Thr Lys Asp His
025                    1030                    1035                    1040
Ala Val Asp Leu Ile Gln Lys Gly Tyr Met Arg Ile Gln Gln Phe Arg
               1045                    1050                    1055
Lys Asn Asp Gly Ser Phe Gly Ala Trp Leu His Arg Asp Ser Ser Thr
           1060                    1065                    1070
Trp Leu Thr Ala Phe Val Leu Lys Ile Leu Ser Leu Ala Gln Glu Gln
       1075                    1080                    1085
Val Gly Asn Ser Pro Glu Lys Leu Gln Glu Thr Ala Ser Trp Leu Leu
   1090                    1095                    1100
Ala Gln Gln Leu Gly Asp Gly Ser Phe His Asp Pro Cys Pro Val Ile
105                    1110                    1115                    1120
His Arg Ala Met Gln Gly Gly Leu Val Gly Ser Asp Glu Thr Val Ala
               1125                    1130                    1135
Leu Thr Ala Phe Val Val Ile Ala Leu His His Gly Leu Asp Val Phe
           1140                    1145                    1150
Gln Asp Asp Asp Ala Lys Gln Leu Lys Asn Arg Val Glu Ala Ser Ile
       1155                    1160                    1165
Thr Lys Ala Asn Ser Phe Leu Gly Gln Lys Ala Ser Ala Gly Leu Leu
   1170                    1175                    1180
Gly Ala His Ala Ala Ala Ile Thr Ala Tyr Ala Leu Thr Leu Thr Lys
185                    1190                    1195                    1200
Ala Ser Glu Asp Leu Arg Asn Val Ala His Asn Ser Leu Met Ala Met
               1205                    1210                    1215
Ala Glu Glu Thr Gly Glu His Leu Tyr Trp Gly Leu Val Leu Gly Ser
           1220                    1225                    1230
Gln Asp Lys Val Val Leu Arg Pro Thr Ala Pro Arg Ser Pro Thr Glu
       1235                    1240                    1245
Pro Val Pro Gln Ala Pro Ala Leu Trp Ile Glu Thr Thr Ala Tyr Ala
   1250                    1255                    1260
Leu Leu His Leu Leu Leu Arg Glu Gly Lys Gly Lys Met Ala Asp Lys
265                    1270                    1275                    1280
Ala Ala Ser Trp Leu Thr His Gln Gly Ser Phe His Gly Ala Phe Arg
               1285                    1290                    1295
Ser Thr Gln Asp Thr Val Val Thr Leu Asp Ala Leu Ser Ala Tyr Trp
           1300                    1305                    1310
Ile Ala Ser His Thr Thr Glu Glu Lys Ala Leu Lys Val Thr Leu Ser
       1315                    1320                    1325
Ser Met Gly Arg Asn Gly Leu Lys Thr His Gly Leu His Leu Asn Asn
   1330                    1335                    1340
His Gln Val Lys Gly Leu Glu Glu Glu Leu Lys Phe Ser Leu Gly Ser
345                    1350                    1355                    1360
Thr Ile Ser Val Lys Val Glu Gly Asn Ser Lys Gly Thr Leu Lys Ile
               1365                    1370                    1375
Leu Arg Thr Tyr Asn Val Leu Asp Met Lys Asn Thr Thr Cys Gln Asp
           1380                    1385                    1390
Leu Gln Ile Glu Val Lys Val Thr Gly Ala Val Glu Tyr Ala Trp Asp
```

1395 1400 1405
Ala Asn Glu Asp Tyr Glu Asp Tyr Tyr Asp Met Pro Ala Ala Asp Asp
1410 1415 1420
Pro Ser Val Pro Leu Gln Pro Val Thr Pro Leu Gln Leu Phe Glu Gly
425 1430 1435 1440
Arg Arg Ser Arg Arg Arg Arg Glu Ala Pro Lys Val Ala Glu Glu Gln
1445 1450 1455
Glu Ser Arg Val Gln Tyr Thr Val Cys Ile Trp Arg Asn Gly Lys Leu
1460 1465 1470
Gly Leu Ser Gly Met Ala Ile Ala Asp Ile Thr Leu Leu Ser Gly Phe
1475 1480 1485
His Ala Leu Arg Ala Asp Leu Glu Lys Leu Thr Ser Leu Ser Asp Arg
1490 1495 1500
Tyr Val Ser His Phe Glu Thr Asp Gly Pro His Val Leu Leu Tyr Phe
505 1510 1515 1520
Asp Ser Val Pro Thr Thr Arg Glu Cys Val Gly Phe Gly Ala Ser Gln
1525 1530 1535
Glu Val Val Val Gly Leu Val Gln Pro Ser Ser Ala Val Leu Tyr Asp
1540 1545 1550
Tyr Tyr Ser Pro Asp His Lys Cys Ser Val Phe Tyr Ala Ala Pro Thr
1555 1560 1565
Lys Ser Gln Leu Leu Ala Thr Leu Cys Ser Gly Asp Val Cys Gln Cys
1570 1575 1580
Ala Glu Gly Lys Cys Pro Arg Leu Leu Arg Ser Leu Glu Arg Arg Val
585 1590 1595 1600
Glu Asp Lys Asp Gly Tyr Arg Met Arg Phe Ala Cys Tyr Tyr Pro Arg
1605 1610 1615
Val Glu Tyr Gly Phe Thr Val Lys Val Leu Arg Glu Asp Gly Arg Ala
1620 1625 1630
Ala Phe Arg Leu Phe Glu Ser Lys Ile Thr Gln Val Leu His Phe Arg
1635 1640 1645
Lys Asp Thr Met Ala Ser Ile Gly Gln Thr Arg Asn Phe Leu Ser Arg
1650 1655 1660
Ala Ser Cys Arg Leu Arg Leu Glu Pro Asn Lys Glu Tyr Leu Ile Met
665 1670 1675 1680
Gly Met Asp Gly Glu Thr Ser Asp Asn Lys Gly Asp Pro Gln Tyr Leu
1685 1690 1695
Leu Asp Ser Asn Thr Trp Ile Glu Glu Met Pro Ser Glu Gln Met Cys
1700 1705 1710
Lys Ser Thr Arg His Arg Ala Ala Cys Phe Gln Leu Lys Asp Phe Leu
1715 1720 1725
Met Glu Phe Ser Ser Arg Gly Cys Gln Val
1730 1735

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1676 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Gly Leu Leu Gly Ile Leu Cys Phe Leu Ile Phe Leu Gly Lys Thr
1 5 10 15
Trp Gly Gln Glu Gln Thr Tyr Val Ile Ser Ala Pro Lys Ile Phe Arg
20 25 30
Val Gly Ala Ser Glu Asn Ile Val Ile Gln Val Tyr Gly Tyr Thr Glu
35 40 45
Ala Phe Asp Ala Thr Ile Ser Ile Lys Ser Tyr Pro Asp Lys Lys Phe
50 55 60
Ser Tyr Ser Ser Gly His Val His Leu Ser Ser Glu Asn Lys Phe Gln
65 70 75 80

```
Asn Ser Ala Ile Leu Thr Ile Gln Pro Lys Gln Leu Pro Gly Gly Gln
                85                  90                  95
Asn Pro Val Ser Tyr Val Tyr Leu Glu Val Val Ser Lys His Phe Ser
            100                 105                 110
Lys Ser Lys Arg Met Pro Ile Thr Tyr Asp Asn Gly Phe Leu Phe Ile
        115                 120                 125
His Thr Asp Lys Pro Val Tyr Thr Pro Asp Gln Ser Val Lys Val Arg
    130                 135                 140
Val Tyr Ser Leu Asn Asp Asp Leu Lys Pro Ala Lys Arg Glu Thr Val
145                 150                 155                 160
Leu Thr Phe Ile Asp Pro Glu Gly Ser Glu Val Asp Met Val Glu Glu
                165                 170                 175
Ile Asp His Ile Gly Ile Ile Ser Phe Pro Asp Phe Lys Ile Pro Ser
            180                 185                 190
Asn Pro Arg Tyr Gly Met Trp Thr Ile Lys Ala Lys Tyr Lys Glu Asp
        195                 200                 205
Phe Ser Thr Thr Gly Thr Ala Tyr Phe Glu Val Lys Glu Tyr Val Leu
    210                 215                 220
Pro His Phe Ser Val Ser Ile Glu Pro Glu Tyr Asn Phe Ile Gly Tyr
225                 230                 235                 240
Lys Asn Phe Lys Asn Phe Glu Ile Thr Ile Lys Ala Arg Tyr Phe Tyr
                245                 250                 255
Asn Lys Val Val Thr Glu Ala Asp Val Tyr Ile Thr Phe Gly Ile Arg
            260                 265                 270
Glu Asp Leu Lys Asp Asp Gln Lys Glu Met Met Gln Thr Ala Met Gln
        275                 280                 285
Asn Thr Met Leu Ile Asn Gly Ile Ala Gln Val Thr Phe Asp Ser Glu
    290                 295                 300
Thr Ala Val Lys Glu Leu Ser Tyr Tyr Ser Leu Glu Asp Leu Asn Asn
305                 310                 315                 320
Lys Tyr Leu Tyr Ile Ala Val Thr Val Ile Glu Ser Thr Gly Gly Phe
                325                 330                 335
Ser Glu Glu Ala Glu Ile Pro Gly Ile Lys Tyr Val Leu Ser Pro Tyr
            340                 345                 350
Lys Leu Asn Leu Val Ala Thr Pro Leu Phe Leu Lys Pro Gly Ile Pro
        355                 360                 365
Tyr Pro Ile Lys Val Gln Val Lys Asp Ser Leu Asp Gln Leu Val Gly
    370                 375                 380
Gly Val Pro Val Ile Leu Asn Ala Gln Thr Ile Asp Val Asn Gln Glu
385                 390                 395                 400
Thr Ser Asp Leu Asp Pro Ser Lys Ser Val Thr Arg Val Asp Asp Gly
                405                 410                 415
Val Ala Ser Phe Val Leu Asn Leu Pro Ser Gly Val Thr Val Leu Glu
            420                 425                 430
Phe Asn Val Lys Thr Asp Ala Pro Asp Leu Pro Glu Glu Asn Gln Ala
        435                 440                 445
Arg Glu Gly Tyr Arg Ala Ile Ala Tyr Ser Ser Leu Ser Gln Ser Tyr
    450                 455                 460
Leu Tyr Ile Asp Trp Thr Asp Asn His Lys Ala Leu Leu Val Gly Glu
465                 470                 475                 480
His Leu Asn Ile Ile Val Thr Pro Lys Ser Pro Tyr Ile Asp Lys Ile
                485                 490                 495
Thr His Tyr Asn Tyr Leu Ile Leu Ser Lys Gly Lys Ile Ile His Phe
            500                 505                 510
Gly Thr Arg Glu Lys Phe Ser Asp Ala Ser Tyr Gln Ser Ile Asn Ile
        515                 520                 525
Pro Val Thr Gln Asn Met Val Pro Ser Ser Arg Leu Leu Val Tyr Tyr
    530                 535                 540
Ile Val Thr Gly Glu Gln Thr Ala Glu Leu Val Ser Asp Ser Val Trp
545                 550                 555                 560
Leu Asn Ile Glu Glu Lys Cys Gly Asn Gln Leu Gln Val His Leu Ser
                565                 570                 575
Pro Asp Ala Asp Ala Tyr Ser Pro Gly Gln Thr Val Ser Leu Asn Met
            580                 585                 590
Ala Thr Gly Met Asp Ser Trp Val Ala Leu Ala Ala Val Asp Ser Ala
        595                 600                 605
Val Tyr Gly Val Gln Arg Gly Ala Lys Lys Pro Leu Glu Arg Val Phe
    610                 615                 620
Gln Phe Leu Glu Lys Ser Asp Leu Gly Cys Gly Ala Gly Gly Gly Leu
625                 630                 635                 640
```

Asn Asn Ala Asn Val Phe His Leu Ala Gly Leu Thr Phe Leu Thr Asn
645 650 655
Ala Asn Ala Asp Asp Ser Gln Glu Asn Asp Glu Pro Cys Lys Glu Ile
660 665 670
Leu Arg Pro Arg Arg Thr Leu Gln Lys Lys Ile Glu Glu Ile Ala Ala
675 680 685
Lys Tyr Lys His Ser Val Val Lys Lys Cys Cys Tyr Asp Gly Ala Cys
690 695 700
Val Asn Asn Asp Glu Thr Cys Glu Gln Arg Ala Ala Arg Ile Ser Leu
705 710 715 720
Gly Pro Arg Cys Ile Lys Ala Phe Thr Glu Cys Cys Val Val Ala Ser
725 730 735
Gln Leu Arg Ala Asn Ile Ser His Lys Asp Met Gln Leu Gly Arg Leu
740 745 750
His Met Lys Thr Leu Leu Pro Val Ser Lys Pro Glu Ile Arg Ser Tyr
755 760 765
Phe Pro Glu Ser Trp Leu Trp Glu Val His Leu Val Pro Arg Arg Lys
770 775 780
Gln Leu Gln Phe Ala Leu Pro Asp Ser Leu Thr Thr Trp Glu Ile Gln
785 790 795 800
Gly Ile Gly Ile Ser Asn Thr Gly Ile Cys Val Ala Asp Thr Val Lys
805 810 815
Ala Lys Val Phe Lys Asp Val Phe Leu Glu Met Asn Ile Pro Tyr Ser
820 825 830
Val Val Arg Gly Glu Gln Ile Gln Leu Lys Gly Thr Val Tyr Asn Tyr
835 840 845
Arg Thr Ser Gly Met Gln Phe Cys Val Lys Met Ser Ala Val Glu Gly
850 855 860
Ile Cys Thr Ser Glu Ser Pro Val Ile Asp His Gln Gly Thr Lys Ser
865 870 875 880
Ser Lys Cys Val Arg Gln Lys Val Glu Gly Ser Ser Ser His Leu Val
885 890 895
Thr Phe Thr Val Leu Pro Leu Glu Ile Gly Leu His Asn Ile Asn Phe
900 905 910
Ser Leu Glu Thr Trp Phe Gly Lys Glu Ile Leu Val Lys Thr Leu Arg
915 920 925
Val Val Pro Glu Gly Val Lys Arg Glu Ser Tyr Ser Gly Val Thr Leu
930 935 940
Asp Pro Arg Gly Ile Tyr Gly Thr Ile Ser Arg Arg Lys Glu Phe Pro
945 950 955 960
Tyr Arg Ile Pro Leu Asp Leu Val Pro Lys Thr Glu Ile Lys Arg Ile
965 970 975
Leu Ser Val Lys Gly Leu Leu Val Gly Glu Ile Leu Ser Ala Val Leu
980 985 990
Ser Gln Glu Gly Ile Asn Ile Leu Thr His Leu Pro Lys Gly Ser Ala
995 1000 1005
Glu Ala Glu Leu Met Ser Val Val Pro Val Phe Tyr Val Phe His Tyr
1010 1015 1020
Leu Glu Thr Gly Asn His Trp Asn Ile Phe His Ser Asp Pro Leu Ile
025 1030 1035 1040
Glu Lys Gln Lys Leu Lys Lys Lys Leu Lys Glu Gly Met Leu Ser Ile
1045 1050 1055
Met Ser Tyr Arg Asn Ala Asp Tyr Ser Tyr Ser Val Trp Lys Gly Gly
1060 1065 1070
Ser Ala Ser Thr Trp Leu Thr Ala Phe Ala Leu Arg Val Leu Gly Gln
1075 1080 1085
Val Asn Lys Tyr Val Glu Gln Asn Gln Asn Ser Ile Cys Asn Ser Leu
1090 1095 1100
Leu Trp Leu Val Glu Asn Tyr Gln Leu Asp Asn Gly Ser Phe Lys Glu
105 1110 1115 1120
Asn Ser Gln Tyr Gln Pro Ile Lys Leu Gln Gly Thr Leu Pro Val Glu
1125 1130 1135
Ala Arg Glu Asn Ser Leu Tyr Leu Thr Ala Phe Thr Val Ile Gly Ile
1140 1145 1150
Arg Lys Ala Phe Asp Ile Cys Pro Leu Val L Ile Asp Thr Ala Leu
1155 1160 1165
Ile Lys Ala Asp Asn Phe Leu Leu Glu Asn Thr Leu Pro Ala Gln Ser
1170 1175 1180
Thr Phe Thr Leu Ala Ile Ser Ala Tyr Ala Leu Ser Leu Gly Asp Lys
185 1190 1195 1200

```
Thr His Pro Gln Phe Arg Ser Ile Val Ser Ala Leu Lys Arg Glu Ala
                1205                1210                1215
Leu Val Lys Gly Asn Pro Pro Ile Tyr Arg Phe Trp Lys Asp Asn Leu
            1220                1225                1230
Gln His Lys Asp Ser Ser Val Pro Asn Thr Gly Thr Ala Arg Met Val
        1235                1240                1245
Glu Thr Thr Ala Tyr Ala Leu Leu Thr Ser Leu Asn Leu Lys Asp Ile
    1250                1255                1260
Asn Tyr Val Asn Pro Val Ile Lys Trp Leu Ser Glu Glu Gln Arg Tyr
265                 1270                1275                1280
Gly Gly Gly Phe Tyr Ser Thr Gln Asp Thr Ile Asn Ala Ile Glu Gly
                1285                1290                1295
Leu Thr Glu Tyr Ser Leu Leu Val Lys Gln Leu Arg Leu Ser Met Asp
            1300                1305                1310
Ile Asp Val Ser Tyr Lys His Lys Gly Ala Leu His Asn Tyr Lys Met
        1315                1320                1325
Thr Asp Lys Asn Phe Leu Gly Arg Pro Val Glu Val Leu Leu Asn Asp
    1330                1335                1340
Asp Leu Ile Val Ser Thr Gly Phe Gly Ser Gly Leu Ala Thr Val His
345                 1350                1355                1360
Val Thr Thr Val Val His Lys Thr Ser Thr Ser Glu Glu Val Cys Ser
                1365                1370                1375
Phe Tyr Leu Lys Ile Asp Thr Gln Asp Ile Glu Ala Ser His Tyr Arg
            1380                1385                1390
Gly Tyr Gly Asn Ser Asp Tyr Lys Arg Ile Val Ala Cys Ala Ser Tyr
        1395                1400                1405
Lys Pro Ser Arg Glu Glu Ser Ser Ser Gly Ser Ser His Ala Val Met
    1410                1415                1420
Asp Ile Ser Leu Pro Thr Gly Ile Ser Ala Asn Glu Glu Asp Leu Lys
425                 1430                1435                1440
Ala Leu Val Glu Gly Val Asp Gln Leu Phe Thr Asp Tyr Gln Ile Lys
                1445                1450                1455
Asp Gly His Val Ile Leu Gln Leu Asn Ser Ile Pro Ser Ser Asp Phe
            1460                1465                1470
Leu Cys Val Arg Phe Arg Ile Phe Glu Leu Phe Glu Val Gly Phe Leu
        1475                1480                1485
Ser Pro Ala Thr Phe Thr Val Tyr Glu Tyr His Arg Pro Asp Lys Gln
    1490                1495                1500
Cys Thr Met Phe Tyr Ser Thr Ser Asn Ile Lys Ile Gln Lys Val Cys
505                 1510                1515                1520
Glu Gly Ala Ala Cys Lys Cys Val Glu Ala Asp Cys Gly Gln Met Gln
                1525                1530                1535
Glu Glu Leu Asp Leu Thr Ile Ser Ala Glu Thr Arg Lys Gln Thr Ala
            1540                1545                1550
Cys Lys Pro Glu Ile Ala Tyr Ala Tyr Lys Val Ser Ile Thr Ser Ile
        1555                1560                1565
Thr Val Glu Asn Val Phe Val Lys Tyr Lys Ala Thr Leu Leu Asp Ile
    1570                1575                1580
Tyr Lys Thr Gly Glu Ala Val Ala Glu Lys Asp Ser Glu Ile Thr Phe
585                 1590                1595                1600
Ile Lys Lys Val Thr Cys Thr Asn Ala Glu Leu Val Lys Gly Arg Gln
                1605                1610                1615
Tyr Leu Ile Met Gly Lys Glu Ala Leu Gln Ile Lys Tyr Asn Phe Ser
            1620                1625                1630
Phe Arg Tyr Ile Tyr Pro Leu Asp Ser Leu Thr Trp Ile Glu Tyr Trp
        1635                1640                1645
Pro Arg Asp Thr Thr Cys Ser Ser Cys Gln Ala Phe Leu Ala Asn Leu
    1650                1655                1660
Asp Glu Phe Ala Glu Asp Ile Phe Leu Asn Gly Cys
665                 1670                1675
```

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1680 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Gly Leu Trp Gly Ile Leu Cys Leu Leu Ile Phe Leu Asp Lys Thr
 1               5                  10                  15
Trp Gly Gln Glu Gln Thr Tyr Val Ile Ser Ala Pro Lys Ile Leu Arg
            20                  25                  30
Val Gly Ser Ser Glu Asn Val Val Ile Gln Val His Gly Tyr Thr Glu
        35                  40                  45
Ala Phe Asp Ala Thr Leu Ser Leu Lys Ser Tyr Pro Asp Lys Lys Val
    50                  55                  60
Thr Phe Ser Ser Gly Tyr Val Asn Leu Ser Pro Glu Asn Lys Phe Gln
65                  70                  75                  80
Asn Ala Ala Leu Leu Thr Leu Gln Pro Asn Gln Val Pro Arg Glu Glu
                85                  90                  95
Ser Pro Val Ser His Val Tyr Leu Glu Val Val Ser Lys His Phe Ser
            100                 105                 110
Lys Ser Lys Lys Ile Pro Ile Thr Tyr Asn Asn Gly Ile Leu Phe Ile
        115                 120                 125
His Thr Asp Lys Pro Val Tyr Thr Pro Asp Gln Ser Val Lys Ile Arg
    130                 135                 140
Val Tyr Ser Leu Gly Asp Asp Leu Lys Pro Ala Lys Arg Glu Thr Val
145                 150                 155                 160
Leu Thr Phe Ile Asp Pro Glu Gly Ser Glu Val Asp Ile Val Glu Glu
                165                 170                 175
Asn Asp Tyr Thr Gly Ile Ile Ser Phe Pro Asp Phe Lys Ile Pro Ser
            180                 185                 190
Asn Pro Lys Tyr Gly Val Trp Thr Ile Lys Ala Asn Tyr Lys Lys Asp
        195                 200                 205
Phe Thr Thr Thr Gly Thr Ala Tyr Phe Glu Ile Lys Glu Tyr Val Leu
    210                 215                 220
Pro Arg Phe Ser Val Ser Ile Glu Leu Glu Arg Thr Phe Ile Gly Tyr
225                 230                 235                 240
Lys Asn Phe Lys Asn Phe Glu Ile Thr Val Lys Ala Arg Tyr Phe Tyr
                245                 250                 255
Asn Lys Val Val Pro Asp Ala Glu Val Tyr Ala Phe Phe Gly Leu Arg
            260                 265                 270
Glu Asp Ile Lys Asp Glu Glu Lys Gln Met Met His Lys Ala Thr Gln
        275                 280                 285
Ala Ala Lys Leu Val Asp Gly Val Ala Gln Ile Ser Phe Asp Ser Glu
    290                 295                 300
Thr Ala Val Lys Glu Leu Ser Tyr Asn Ser Leu Glu Asp Leu Asn Asn
305                 310                 315                 320
Lys Tyr Leu Tyr Ile Ala Val Thr Val Thr Glu Ser Ser Gly Gly Phe
                325                 330                 335
Ser Glu Glu Ala Glu Ile Pro Gly Val Lys Tyr Val Leu Ser Pro Tyr
            340                 345                 350
Thr Leu Asn Leu Val Ala Thr Pro Leu Phe Val Lys Pro Gly Ile Pro
        355                 360                 365
Phe Ser Ile Lys Ala Gln Val Lys Asp Ser Leu Glu Gln Ala Val Gly
    370                 375                 380
Gly Val Pro Val Thr Leu Met Ala Gln Thr Val Asp Val Asn Gln Glu
385                 390                 395                 400
Thr Ser Asp Leu Glu Thr Lys Arg Ser Ile Thr His Asp Thr Asp Gly
                405                 410                 415
Val Ala Val Phe Val Leu Asn Leu Pro Ser Asn Val Thr Val Leu Lys
            420                 425                 430
Phe Glu Ile Arg Thr Asp Asp Pro Glu Leu Pro Glu Glu Asn Gln Ala
        435                 440                 445
Ser Lys Glu Tyr Glu Ala Val Ala Tyr Ser Ser Leu Ser Gln Ser Tyr
    450                 455                 460
Ile Tyr Ile Ala Trp Thr Glu Asn Tyr Lys Pro Met Leu Val Gly Glu
465                 470                 475                 480
Tyr Leu Asn Ile Met Val Thr Pro Lys Ser Pro Tyr Ile Asp Lys Ile
                485                 490                 495
Thr His Tyr Asn Tyr Leu Ile Leu Ser Lys Gly Lys Ile Val Gln Tyr
```

500 505 510
Gly Thr Arg Glu Lys Leu Phe Ser Ser Thr Tyr Gln Asn Ile Asn Ile
515 520 525
Pro Val Thr Gln Asn Met Val Pro Ser Ala Arg Leu Leu Val Tyr Tyr
530 535 540
Ile Val Thr Gly Glu Gln Thr Ala Glu Leu Val Ala Asp Ala Val Trp
545 550 555 560
Ile Asn Ile Glu Glu Lys Cys Gly Asn Gln Leu Gln Val His Leu Ser
565 570 575
Pro Asp Glu Tyr Val Tyr Ser Pro Gly Gln Thr Val Ser Leu Asp Met
580 585 590
Val Thr Glu Ala Asp Ser Trp Val Ala Leu Ser Ala Val Asp Arg Ala
595 600 605
Val Tyr Lys Val Gln Gly Asn Ala Lys Arg Ala Met Gln Arg Val Phe
610 615 620
Gln Ala Leu Asp Glu Lys Ser Asp Leu Gly Cys Gly Ala Gly Gly Gly
625 630 635 640
His Asp Asn Ala Asp Val Phe His Leu Ala Gly Leu Thr Phe Leu Thr
645 650 655
Asn Ala Asn Ala Asp Asp Ser His Tyr Arg Asp Asp Ser Cys Lys Glu
660 665 670
Ile Leu Arg Ser Lys Arg Asn Leu His Leu Leu Arg Gln Lys Ile Glu
675 680 685
Glu Gln Ala Ala Lys Tyr Lys His Ser Val Pro Lys Lys Cys Cys Tyr
690 695 700
Asp Gly Ala Arg Val Asn Phe Tyr Glu Thr Cys Glu Glu Arg Val Ala
705 710 715 720
Arg Val Thr Ile Gly Pro Leu Cys Ile Arg Ala Phe Asn Glu Cys Cys
725 730 735
Thr Ile Ala Asn Lys Ile Arg Lys Glu Ser Pro His Lys Pro Val Gln
740 745 750
Leu Gly Arg Ile His Ile Lys Thr Leu Leu Pro Val Met Lys Ala Asp
755 760 765
Ile Arg Ser Tyr Phe Pro Glu Ser Trp Leu Trp Glu Ile His Arg Val
770 775 780
Pro Lys Arg Lys Gln Leu Gln Val Thr Leu Pro Asp Ser Leu Thr Thr
785 790 795 800
Trp Glu Ile Gln Gly Ile Gly Ile Ser Asp Asn Gly Ile Cys Val Ala
805 810 815
Asp Thr Leu Lys Ala Lys Val Phe Lys Glu Val Phe Leu Glu Met Asn
820 825 830
Ile Pro Tyr Ser Val Val Arg Gly Glu Gln Ile Gln Leu Lys Gly Thr
835 840 845
Val Tyr Asn Tyr Met Thr Ser Gly Thr Lys Phe Cys Val Lys Met Ser
850 855 860
Ala Val Glu Gly Ile Cys Thr Ser Gly Ser Ser Ala Ala Ser Leu His
865 870 875 880
Thr Ser Arg Pro Ser Arg Cys Val Phe Gln Arg Ile Glu Gly Ser Ser
885 890 895
Ser His Leu Val Thr Phe Thr Leu Leu Pro Leu Glu Ile Gly Leu His
900 905 910
Ser Ile Asn Phe Ser Leu Glu Thr Ser Phe Gly Lys Asp Ile Leu Val
915 920 925
Lys Thr Leu Arg Val Val Pro Glu Gly Val Lys Arg Glu Ser Tyr Ala
930 935 940
Gly Val Ile Leu Asp Pro Lys Gly Ile Arg Gly Ile Val Asn Arg Arg
945 950 955 960
Lys Glu Phe Pro Tyr Arg Ile Pro Leu Asp Leu Val Pro Lys Thr Lys
965 970 975
Val Glu Arg Ile Leu Ser Val Lys Gly Leu Leu Val Gly Glu Phe Leu
980 985 990
Ser Thr Val Leu Ser Lys Glu Gly Ile Asn Ile Leu Thr His Leu Pro
995 1000 1005
Lys Gly Ser Ala Glu Ala Glu Leu Met Ser Ile Ala Pro Val Phe Tyr
1010 1015 1020
Val Phe His Tyr Leu Glu Ala Gly Asn His Trp Asn Ile Phe Tyr Pro
025 1030 1035 1040
Asp Thr Leu Ser Lys Arg Gln Ser Leu Glu Lys Lys Ile Lys Gln Gly
1045 1050 1055
Val Val Ser Val Met Ser Tyr Arg Asn Ala Asp Tyr Ser Tyr Ser Met

1060 1065 1070
Trp Lys Gly Ala Ser Ala Ser Thr Trp Leu Thr Ala Phe Ala Leu Arg
1075 1080 1085
Val Leu Gly Gln Val Ala Lys Tyr Val Lys Gln Asp Glu Asn Ser Ile
1090 1095 1100
Cys Asn Ser Leu Leu Trp Leu Val Glu Lys Cys Gln Leu Glu Asn Gly
105 1110 1115 1120
Ser Phe Lys Glu Asn Ser Gln Tyr Leu Pro Ile Lys Leu Gln Gly Thr
1125 1130 1135
Leu Pro Ala Glu Ala Gln Glu Lys Thr Leu Tyr Leu Thr Ala Phe Ser
1140 1145 1150
Val Ile Gly Ile Arg Lys Ala Val Asp Ile Cys Pro Thr Met Lys Ile
1155 1160 1165
His Thr Ala Leu Asp Lys Ala Asp Ser Phe Leu Leu Glu Asn Thr Leu
1170 1175 1180
Pro Ser Lys Ser Thr Phe Thr Leu Ala Ile Val Ala Tyr Ala Leu Ser
185 1190 1195 1200
Leu Gly Asp Arg Thr His Pro Arg Phe Arg Leu Ile Val Ser Ala Leu
1205 1210 1215
Arg Lys Glu Ala Phe Val Lys Gly Asp Pro Pro Ile Tyr Arg Tyr Trp
1220 1225 1230
Arg Asp Thr Leu Lys Arg Pro Asp Ser Ser Val Pro Ser Ser Gly Thr
1235 1240 1245
Ala Gly Met Val Glu Thr Thr Ala Tyr Ala Leu Leu Ala Ser Leu Lys
1250 1255 1260
Leu Lys Asp Met Asn Tyr Ala Asn Pro Ile Ile Lys Trp Leu Ser Glu
265 1270 1275 1280
Glu Gln Arg Tyr Gly Gly Gly Phe Tyr Ser Thr Gln Asp Thr Ile Asn
1285 1290 1295
Ala Ile Glu Gly Leu Thr Glu Tyr Ser Leu Leu Leu Lys Gln Ile His
1300 1305 1310
Leu Asp Met Asp Ile Asn Val Ala Tyr Lys His Glu Gly Asp Phe His
1315 1320 1325
Lys Tyr Lys Val Thr Glu Lys His Phe Leu Gly Arg Pro Val Glu Val
1330 1335 1340
Ser Leu Asn Asp Asp Leu Val Val Ser Thr Gly Tyr Ser Ser Gly Leu
345 1350 1355 1360
Ala Thr Val Tyr Val Lys Thr Val Val His Lys Ile Ser Val Ser Glu
1365 1370 1375
Glu Phe Cys Ser Phe Tyr Leu Lys Ile Asp Thr Gln Asp Ile Glu Ala
1380 1385 1390
Ser Ser His Phe Arg Leu Ser Asp Ser Gly Phe Lys Arg Ile Ile Ala
1395 1400 1405
Cys Ala Ser Tyr Lys Pro Ser Lys Glu Glu Ser Thr Ser Gly Ser Ser
1410 1415 1420
His Ala Val Met Asp Ile Ser Leu Pro Thr Gly Ile Gly Ala Asn Glu
425 1430 1435 1440
Glu Asp Leu Arg Ala Leu Val Glu Gly Val Asp Gln Leu Leu Thr Asp
1445 1450 1455
Tyr Gln Ile Lys Asp Gly His Val Ile Leu Gln Leu Asn Ser Ile Pro
1460 1465 1470
Ser Arg Asp Phe Leu Cys Val Arg Phe Arg Ile Phe Glu Leu Phe Gln
1475 1480 1485
Val Gly Phe Leu Asn Pro Ala Thr Phe Thr Val Tyr Glu Tyr His Arg
1490 1495 1500
Pro Asp Lys Gln Cys Thr Met Ile Tyr Ser Ile Ser Asp Thr Arg Leu
505 1510 1515 1520
Gln Lys Val Cys Glu Gly Ala Ala Cys Thr Cys Val Glu Ala Asp Cys
1525 1530 1535
Ala Gln Leu Gln Ala Glu Val Asp Leu Ala Ile Ser Ala Asp Ser Arg
1540 1545 1550
Lys Glu Lys Ala Cys Lys Pro Glu Thr Ala Tyr Ala Tyr Lys Val Arg
1555 1560 1565
Ile Thr Ser Ala Thr Glu Glu Asn Val Phe Val Lys Tyr Thr Ala Thr
1570 1575 1580
Leu Leu Val Thr Tyr Lys Thr Gly Glu Ala Ala Asp Glu Asn Ser Glu
585 1590 1595 1600
Val Thr Phe Ile Lys Lys Met Ser Cys Thr Asn Ala Asn Leu Val Lys
1605 1610 1615
Gly Lys Gln Tyr Leu Ile Met Gly Lys Glu Val Leu Gln Ile Lys His

```
            1620                    1625                    1630
Asn Phe Ser Phe Lys Tyr Ile Tyr Pro Leu Asp Ser Ser Thr Trp Ile
        1635                    1640                    1645
Glu Tyr Trp Pro Thr Asp Thr Thr Cys Pro Ser Cys Gln Ala Phe Val
    1650                    1655                    1660
Glu Asn Leu Asn Asn Phe Ala Glu Asp Leu Phe Leu Asn Ser Cys Glu
665                 1670                    1675                    1680
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1235 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
Met Phe Ser Gly Gly Gly Gly Pro Leu Ser Pro Gly Gly Lys Ser Ala
 1               5                  10                  15
Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro Ala Gly Pro Arg Gly Ala
            20                  25                  30
Gly Arg Gly Pro Pro Pro Cys Leu Arg Gln Asn Phe Tyr Asn Pro Tyr
        35                  40                  45
Leu Ala Pro Val Gly Thr Gln Gln Lys Pro Thr Gly Pro Thr Gln Arg
    50                  55                  60
His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro Arg
65                  70                  75                  80
Val Leu Asp Glu Asp Ala Pro Pro Glu Lys Arg Ala Gly Val His Asp
                85                  90                  95
Gly His Leu Lys Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu Arg
            100                 105                 110
Asp Val Leu Arg Val Gly Ser Gly Gly Phe Trp Pro Arg Arg Ser Arg
        115                 120                 125
Leu Trp Gly Gly Val Asp His Ala Pro Ala Gly Phe Asn Pro Thr Val
    130                 135                 140
Thr Val Phe His Val Tyr Asp Ile Leu Glu Asn Val Glu His Ala Tyr
145                 150                 155                 160
Gly Met Arg Ala Ala Gln Phe His Ala Arg Phe Met Asp Ala Ile Thr
                165                 170                 175
Pro Thr Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly His
            180                 185                 190
Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met Asn
        195                 200                 205
Lys Glu Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp Leu
    210                 215                 220
Cys Glu Arg Met Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser Phe
225                 230                 235                 240
Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg Thr
                245                 250                 255
Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Ala Leu Phe Tyr Arg Val Tyr
            260                 265                 270
Val Arg Ser Gly Arg Val Leu Ser Tyr Leu Cys Asp Asn Phe Cys Pro
        275                 280                 285
Ala Ile Lys Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe Ile
    290                 295                 300
Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys Pro
305                 310                 315                 320
Gly Arg Asn Asn Thr Leu Ala Gln Pro Arg Ala Pro Met Ala Phe Gly
                325                 330                 335
Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala Ile
            340                 345                 350
Glu Gly Gly Met Ser Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe Asp
        355                 360                 365
```

Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val Ala
370 375 380
Gly His Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr Asp
385 390 395 400
Leu Ser Thr Thr Ala Leu Glu His Val Leu Leu Phe Ser Leu Gly Ser
405 410 415
Cys Asp Leu Pro Glu Ser His Leu Asn Glu Leu Ala Ala Arg Gly Leu
420 425 430
Pro Thr Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu Leu
435 440 445
Ala Phe Met Thr Leu Val Lys Gln Tyr Gly Pro Glu Phe Val Thr Gly
450 455 460
Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Leu Leu Ala Lys Leu Thr
465 470 475 480
Asp Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly Arg
485 490 495
Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys Arg
500 505 510
Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly Ile
515 520 525
Ile Thr Asp Lys Ile Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val Ala
530 535 540
Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp Ile
545 550 555 560
Pro Ala Tyr Tyr Ala Thr Gly Pro Ala Gln Arg Gly Val Ile Gly Glu
565 570 575
Tyr Cys Ile Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys Phe
580 585 590
Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile Asn
595 600 605
Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr Cys
610 615 620
Leu Leu Arg Leu Ala Asp Gln Lys Gly Phe Ile Leu Pro Asp Thr Gln
625 630 635 640
Gly Arg Phe Arg Gly Ala Gly Gly Glu Ala Pro Lys Arg Pro Ala Ala
645 650 655
Ala Arg Glu Asp Glu Glu Arg Pro Glu Glu Glu Gly Glu Asp Glu Asp
660 665 670
Glu Arg Glu Glu Gly Gly Gly Glu Arg Glu Pro Glu Gly Ala Arg Glu
675 680 685
Thr Ala Gly Arg His Val Gly Tyr Gln Gly Ala Lys Val Leu Asp Pro
690 695 700
Thr Ser Gly Phe His Val Asn Pro Val Val Val Phe Asp Phe Ala Ser
705 710 715 720
Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu Cys Phe Ser Thr Leu
725 730 735
Ser Leu Arg Ala Asp Ala Val Ala His Leu Glu Ala Gly Lys Asp Tyr
740 745 750
Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe Phe Val Lys Ala His
755 760 765
Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg Asp Trp Leu Ala Met
770 775 780
Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser Ser Pro Glu Glu Ala
785 790 795 800
Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys Val Val Cys Asn Ser
805 810 815
Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu Leu Pro Cys Leu His
820 825 830
Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu Met Leu Leu Ala Thr
835 840 845
Arg Glu Tyr Val His Ala Arg Trp Ala Ala Phe Glu Gln Leu Leu Ala
850 855 860
Asp Phe Pro Glu Ala Ala Asp Met Arg Ala Pro Gly Pro Tyr Ser Met
865 870 875 880
Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe Val Leu Cys Arg Gly
885 890 895
Leu Thr Ala Ala Gly Leu Thr Ala Met Gly Asp Lys Met Ala Ser His
900 905 910
Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys Leu Glu Cys Glu Lys
915 920 925

Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys Lys Tyr Ile Gly Val
930 935 940
Ile Tyr Gly Gly Lys Met Leu Ile Lys Gly Val Asp Leu Val Arg Lys
945 950 955 960
Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg Ala Leu Val Asp Leu
965 970 975
Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala Ala Ala Leu Ala Glu
980 985 990
Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu Pro Glu Gly Leu Gln
995 1000 1005
Ala Phe Gly Ala Val Leu Val Asp Ala His Arg Arg Ile Thr Asp Pro
1010 1015 1020
Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala Glu Leu Ser Arg His
025 1030 1035 1040
Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His Leu Thr Val Tyr Tyr
1045 1050 1055
Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser Ile Lys Asp Arg Ile
1060 1065 1070
Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val Glu Glu Thr Val Ala
1075 1080 1085
Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala Ala Ala Pro Gly Asp Glu
1090 1095 1100
Pro Ala Pro Pro Ala Ala Leu Pro Ser Pro Ala Lys Arg Pro Arg Glu
105 1110 1115 1120
Thr Pro Ser His Ala Asp Pro Pro Gly Gly Ala Ser Lys Pro Arg Lys
1125 1130 1135
Leu Leu Val Ser Glu Leu Ala Glu Asp Pro Ala Tyr Ala Ile Ala His
1140 1145 1150
Gly Val Ala Leu Asn Thr Asp Tyr Tyr Phe Ser His Leu Leu Gly Ala
1155 1160 1165
Ala Cys Val Thr Phe Lys Ala Leu Phe Gly Asn Asn Ala Lys Ile Thr
1170 1175 1180
Glu Ser Leu Leu Lys Arg Phe Ile Pro Glu Val Trp His Pro Pro Asp
185 1190 1195 1200
Asp Val Ala Ala Arg Leu Arg Ala Ala Gly Phe Gly Ala Val Gly Ala
1205 1210 1215
Gly Ala Thr Ala Glu Glu Thr Arg Arg Met Leu His Arg Ala Phe Asp
1220 1225 1230
Thr Leu Ala
1235

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1240 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Phe Cys Ala Ala Gly Gly Pro Ala Ser Pro Gly Gly Lys Ser Ala
1 5 10 15
Ala Arg Ala Ala Ser Gly Phe Phe Ala Pro His Asn Pro Arg Gly Ala
20 25 30
Thr Gln Thr Ala Pro Pro Pro Cys Arg Arg Gln Asn Phe Tyr Asn Pro
35 40 45
His Leu Ala Gln Thr Gly Thr Gln Pro Lys Ala Pro Gly Pro Ala Gln
50 55 60
Arg His Thr Tyr Tyr Ser Glu Cys Asp Glu Phe Arg Phe Ile Ala Pro
65 70 75 80
Arg Ser Leu Asp Glu Asp Ala Pro Ala Glu Gln Arg Thr Gly Val His
85 90 95
Asp Gly Arg Leu Arg Arg Ala Pro Lys Val Tyr Cys Gly Gly Asp Glu

100 105 110
Arg Asp Val Leu Arg Val Gly Pro Glu Gly Phe Trp Pro Arg Arg Leu
115 120 125
Arg Leu Trp Gly Gly Ala Asp His Ala Pro Glu Gly Phe Asp Pro Thr
130 135 140
Val Thr Val Phe His Val Tyr Asp Ile Leu Glu His Val Glu His Ala
145 150 155 160
Tyr Ser Met Arg Ala Ala Gln Leu His Glu Arg Phe Met Asp Ala Ile
165 170 175
Thr Pro Ala Gly Thr Val Ile Thr Leu Leu Gly Leu Thr Pro Glu Gly
180 185 190
His Arg Val Ala Val His Val Tyr Gly Thr Arg Gln Tyr Phe Tyr Met
195 200 205
Asn Lys Ala Glu Val Asp Arg His Leu Gln Cys Arg Ala Pro Arg Asp
210 215 220
Leu Cys Glu Arg Leu Ala Ala Ala Leu Arg Glu Ser Pro Gly Ala Ser
225 230 235 240
Phe Arg Gly Ile Ser Ala Asp His Phe Glu Ala Glu Val Val Glu Arg
245 250 255
Ala Asp Val Tyr Tyr Tyr Glu Thr Arg Pro Thr Leu Tyr Tyr Arg Val
260 265 270
Phe Val Arg Ser Gly Arg Ala Leu Ala Tyr Leu Cys Asp Asn Phe Cys
275 280 285
Pro Ala Ile Arg Lys Tyr Glu Gly Gly Val Asp Ala Thr Thr Arg Phe
290 295 300
Ile Leu Asp Asn Pro Gly Phe Val Thr Phe Gly Trp Tyr Arg Leu Lys
305 310 315 320
Pro Gly Arg Gly Asn Ala Pro Ala Gln Pro Arg Pro Pro Thr Ala Phe
325 330 335
Gly Thr Ser Ser Asp Val Glu Phe Asn Cys Thr Ala Asp Asn Leu Ala
340 345 350
Val Glu Gly Ala Met Cys Asp Leu Pro Ala Tyr Lys Leu Met Cys Phe
355 360 365
Asp Ile Glu Cys Lys Ala Gly Gly Glu Asp Glu Leu Ala Phe Pro Val
370 375 380
Ala Glu Arg Pro Glu Asp Leu Val Ile Gln Ile Ser Cys Leu Leu Tyr
385 390 395 400
Asp Leu Ser Thr Thr Ala Leu Glu His Ile Leu Leu Phe Ser Leu Gly
405 410 415
Ser Cys Asp Leu Pro Glu Ser His Leu Ser Asp Leu Ala Ser Arg Gly
420 425 430
Leu Pro Ala Pro Val Val Leu Glu Phe Asp Ser Glu Phe Glu Met Leu
435 440 445
Leu Ala Phe Met Thr Phe Val Lys Gln Tyr Gly Pro Glu Phe Val Thr
450 455 460
Gly Tyr Asn Ile Ile Asn Phe Asp Trp Pro Phe Val Leu Thr Lys Leu
465 470 475 480
Thr Glu Ile Tyr Lys Val Pro Leu Asp Gly Tyr Gly Arg Met Asn Gly
485 490 495
Arg Gly Val Phe Arg Val Trp Asp Ile Gly Gln Ser His Phe Gln Lys
500 505 510
Arg Ser Lys Ile Lys Val Asn Gly Met Val Asn Ile Asp Met Tyr Gly
515 520 525
Ile Ile Thr Asp Lys Val Lys Leu Ser Ser Tyr Lys Leu Asn Ala Val
530 535 540
Ala Glu Ala Val Leu Lys Asp Lys Lys Lys Asp Leu Ser Tyr Arg Asp
545 550 555 560
Ile Pro Ala Tyr Tyr Ala Ser Gly Pro Ala Gln Arg Gly Val Ile Gly
565 570 575
Glu Tyr Cys Val Gln Asp Ser Leu Leu Val Gly Gln Leu Phe Phe Lys
580 585 590
Phe Leu Pro His Leu Glu Leu Ser Ala Val Ala Arg Leu Ala Gly Ile
595 600 605
Asn Ile Thr Arg Thr Ile Tyr Asp Gly Gln Gln Ile Arg Val Phe Thr
610 615 620
Cys Leu Leu Arg Leu Ala Gly Gln Lys Gly Phe Ile Leu Pro Asp Thr
625 630 635 640
Gln Gly Arg Phe Arg Gly Leu Asp Lys Glu Ala Pro Lys Arg Pro Ala
645 650 655
Val Pro Arg Gly Glu Gly Glu Arg Pro Gly Asp Gly Asn Gly Asp Glu

```
                  660                      665                      670
Asp Lys Asp Asp Asp Glu Asp Gly Asp Glu Asp Gly Asp Glu Arg Glu
            675                      680                      685
Glu Val Ala Arg Glu Thr Gly Gly Arg His Val Gly Tyr Gln Gly Ala
        690                      695                      700
Arg Val Leu Asp Pro Thr Ser Gly Phe His Val Asp Pro Val Val Val
705                     710                     715                     720
Phe Asp Phe Ala Ser Leu Tyr Pro Ser Ile Ile Gln Ala His Asn Leu
                725                     730                     735
Cys Phe Ser Thr Leu Ser Leu Arg Pro Glu Ala Val Ala His Leu Glu
            740                     745                     750
Ala Asp Arg Asp Tyr Leu Glu Ile Glu Val Gly Gly Arg Arg Leu Phe
        755                     760                     765
Phe Val Lys Ala His Val Arg Glu Ser Leu Leu Ser Ile Leu Leu Arg
    770                     775                     780
Asp Trp Leu Ala Met Arg Lys Gln Ile Arg Ser Arg Ile Pro Gln Ser
785                     790                     795                     800
Pro Pro Glu Glu Ala Val Leu Leu Asp Lys Gln Gln Ala Ala Ile Lys
                805                     810                     815
Val Val Cys Asn Ser Val Tyr Gly Phe Thr Gly Val Gln His Gly Leu
            820                     825                     830
Leu Pro Cys Leu His Val Ala Ala Thr Val Thr Thr Ile Gly Arg Glu
        835                     840                     845
Met Leu Leu Ala Thr Arg Ala Tyr Val His Ala Arg Trp Ala Glu Phe
    850                     855                     860
Asp Gln Leu Leu Ala Asp Phe Pro Glu Ala Ala Gly Met Arg Ala Pro
865                     870                     875                     880
Gly Pro Tyr Ser Met Arg Ile Ile Tyr Gly Asp Thr Asp Ser Ile Phe
                885                     890                     895
Val Leu Cys Arg Gly Leu Thr Gly Glu Ala Leu Val Ala Met Gly Asp
            900                     905                     910
Lys Met Ala Ser His Ile Ser Arg Ala Leu Phe Leu Pro Pro Ile Lys
        915                     920                     925
Leu Glu Cys Glu Lys Thr Phe Thr Lys Leu Leu Leu Ile Ala Lys Lys
    930                     935                     940
Lys Tyr Ile Gly Val Ile Cys Gly Gly Lys Met Leu Ile Lys Gly Val
945                     950                     955                     960
Asp Leu Val Arg Lys Asn Asn Cys Ala Phe Ile Asn Arg Thr Ser Arg
                965                     970                     975
Ala Leu Val Asp Leu Leu Phe Tyr Asp Asp Thr Val Ser Gly Ala Ala
            980                     985                     990
Ala Ala Leu Ala Glu Arg Pro Ala Glu Glu Trp Leu Ala Arg Pro Leu
        995                     1000                    1005
Pro Glu Gly Leu Gln Ala Phe Gly Ala Val Leu Val Asp Ala His Arg
    1010                    1015                    1020
Arg Ile Thr Asp Pro Glu Arg Asp Ile Gln Asp Phe Val Leu Thr Ala
025                     1030                    1035                    1040
Glu Leu Ser Arg His Pro Arg Ala Tyr Thr Asn Lys Arg Leu Ala His
                1045                    1050                    1055
Leu Thr Val Tyr Tyr Lys Leu Met Ala Arg Arg Ala Gln Val Pro Ser
            1060                    1065                    1070
Ile Lys Asp Arg Ile Pro Tyr Val Ile Val Ala Gln Thr Arg Glu Val
        1075                    1080                    1085
Glu Glu Thr Val Ala Arg Leu Ala Ala Leu Arg Glu Leu Asp Ala Ala
    1090                    1095                    1100
Ala Pro Gly Asp Glu Pro Ala Pro Pro Ala Ala Leu Pro Ser Pro Ala
105                     1110                    1115                    1120
Lys Arg Pro Arg Glu Thr Pro Ser His Ala Asp Pro Pro Gly Gly Ala
                1125                    1130                    1135
Ser Lys Pro Arg Lys Leu Leu Val Ser Glu Leu Ala Glu Asp Pro Gly
            1140                    1145                    1150
Tyr Ala Ile Ala Arg Gly Val Pro Leu Asn Thr Asp Tyr Tyr Phe Ser
        1155                    1160                    1165
His Leu Leu Gly Ala Ala Cys Val Thr Phe Lys Ala Leu Phe Gly Asn
    1170                    1175                    1180
Asn Ala Lys Ile Thr Glu Ser Leu Leu Lys Arg Phe Ile Pro Glu Thr
185                     1190                    1195                    1200
Trp His Pro Pro Asp Asp Val Ala Ala Arg Leu Arg Ala Ala Gly Phe
                1205                    1210                    1215
Gly Pro Ala Gly Ala Gly Ala Thr Ala Glu Glu Thr Arg Arg Met Leu
```

```
            1220                     1225                     1230
His Arg Ala Phe Asp Thr Leu Ala
        1235                 1240
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1012 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
Met Asp Ser Val Ser Phe Phe Asn Pro Tyr Leu Glu Ala Asn Arg Leu
1               5                   10                  15
Lys Lys Lys Ser Arg Ser Ser Tyr Ile Arg Ile Leu Pro Arg Gly Ile
            20                  25                  30
Met His Asp Gly Ala Ala Gly Leu Ile Lys Asp Val Cys Asp Ser Glu
        35                  40                  45
Pro Arg Met Phe Tyr Arg Asp Arg Gln Tyr Leu Leu Ser Lys Glu Met
    50                  55                  60
Thr Trp Pro Ser Leu Asp Ile Ala Arg Ser Lys Asp Tyr Asp His Met
65                  70                  75                  80
Arg Met Lys Phe His Ile Tyr Asp Ala Val Glu Thr Leu Met Phe Thr
                85                  90                  95
Asp Ser Ile Glu Asn Leu Pro Phe Gln Tyr Arg His Phe Val Ile Pro
            100                 105                 110
Ser Gly Thr Val Ile Arg Met Phe Gly Arg Thr Glu Asp Gly Glu Lys
        115                 120                 125
Ile Cys Val Asn Val Phe Gly Gln Glu Gln Tyr Phe Tyr Cys Glu Cys
    130                 135                 140
Val Asp Gly Arg Ser Leu Lys Ala Thr Ile Asn Asn Leu Met Leu Thr
145                 150                 155                 160
Gly Glu Val Lys Met Ser Cys Ser Phe Val Ile Glu Pro Ala Asp Lys
                165                 170                 175
Leu Ser Leu Tyr Gly Tyr Asn Ala Asn Thr Val Val Asn Leu Phe Lys
            180                 185                 190
Val Ser Phe Gly Asn Phe Tyr Val Ser Gln Arg Ile Gly Lys Ile Leu
        195                 200                 205
Gln Asn Glu Gly Phe Val Val Tyr Glu Ile Asp Val Asp Val Leu Thr
    210                 215                 220
Arg Phe Phe Val Asp Asn Gly Phe Leu Ser Phe Gly Trp Tyr Asn Val
225                 230                 235                 240
Lys Lys Tyr Ile Pro Gln Asp Met Gly Lys Gly Ser Asn Leu Glu Val
                245                 250                 255
Glu Ile Asn Cys His Val Ser Asp Leu Val Ser Leu Glu Asp Val Asn
            260                 265                 270
Trp Pro Leu Tyr Gly Cys Trp Ser Phe Asp Ile Glu Cys Leu Gly Gln
        275                 280                 285
Asn Gly Asn Phe Pro Asp Ala Glu Asn Leu Gly Asp Ile Val Ile Gln
    290                 295                 300
Ile Ser Val Ile Ser Phe Asp Thr Glu Gly Asp Arg Asp Glu Arg His
305                 310                 315                 320
Leu Phe Thr Leu Gly Thr Cys Glu Lys Ile Asp Gly Val His Ile Tyr
                325                 330                 335
Glu Phe Ala Ser Glu Phe Glu Leu Leu Leu Gly Phe Phe Ile Phe Leu
            340                 345                 350
Arg Ile Glu Ser Pro Glu Phe Ile Thr Gly Ty  Asn Ile Asn Asn Phe
        355                 360                 365
Asp Leu Lys Tyr Leu Cys Ile Arg Met Asp Lys Ile Tyr His Tyr Asp
    370                 375                 380
Ile Gly Cys Phe Ser Lys Leu Lys Asn Gly Lys Ile Gly Ile Ser Val
385                 390                 395                 400
```

Pro His Glu Gln Tyr Arg Lys Gly Phe Leu Gln Ala Gln Thr Lys Val
405 410 415
Phe Thr Ser Gly Val Leu Tyr Leu Asp Met Tyr Pro Val Tyr Ser Ser
420 425 430
Lys Ile Thr Ala Gln Asn Tyr Lys Leu Asp Thr Ile Ala Lys Ile Cys
435 440 445
Leu Gln Gln Glu Lys Glu Gln Leu Ser Tyr Lys Glu Ile Pro Lys Lys
450 455 460
Phe Ile Ser Gly Pro Ser Gly Arg Ala Val Val Gly Lys Tyr Cys Leu
465 470 475 480
Gln Asp Ser Val Leu Val Val Arg Leu Phe Lys Gln Ile Asn Tyr His
485 490 495
Phe Glu Val Ala Glu Val Ala Arg Leu Ala His Val Thr Ala Arg Cys
500 505 510
Val Val Phe Glu Gly Gln Gln Lys Lys Ile Phe Pro Cys Ile Leu Thr
515 520 525
Glu Ala Lys Arg Arg Asn Met Ile Leu Pro Ser Met Val Ser Ser His
530 535 540
Asn Arg Gln Gly Ile Gly Tyr Lys Gly Ala Thr Val Leu Glu Pro Lys
545 550 555 560
Thr Gly Tyr Tyr Ala Val Pro Thr Val Val Phe Asp Phe Gln Ser Leu
565 570 575
Tyr Pro Ser Ile Met Met Ala His Asn Leu Cys Tyr Ser Thr Leu Val
580 585 590
Leu Asp Glu Arg Gln Ile Ala Gly Leu Ser Glu Ser Asp Ile Leu Thr
595 600 605
Val Lys Leu Gly Asp Glu Thr His Arg Phe Val Lys Pro Cys Ile Arg
610 615 620
Glu Ser Val Leu Gly Ser Leu Leu Lys Asp Trp Leu Ala Lys Arg Arg
625 630 635 640
Glu Val Lys Ala Glu Met Gln Asn Cys Ser Asp Pro Met Met Lys Leu
645 650 655
Leu Leu Asp Lys Lys Gln Leu Ala Leu Lys Thr Thr Cys Asn Ser Val
660 665 670
Tyr Gly Val Thr Gly Ala Ala His Gly Leu Leu Pro Cys Val Ala Ile
675 680 685
Ala Ala Ser Val Thr Cys Leu Gly Arg Glu Met Leu Cys Ser Thr Val
690 695 700
Asp Tyr Val Asn Ser Lys Met Gln Ser Glu Gln Phe Phe Cys Glu Glu
705 710 715 720
Phe Gly Leu Thr Ser Ser Asp Phe Thr Gly Asp Leu Glu Val Glu Val
725 730 735
Ile Tyr Gly Asp Thr Asp Ser Ile Phe Met Ser Val Arg Asn Met Val
740 745 750
Asn Gln Ser Leu Arg Arg Ile Ala Pro Met Ile Ala Lys His Ile Thr
755 760 765
Asp Arg Leu Phe Lys Ser Pro Ile Lys Leu Glu Phe Glu Lys Ile Leu
770 775 780
Cys Pro Leu Ile Leu Ile Cys Lys Lys Arg Tyr Ile Gly Arg Gln Asp
785 790 795 800
Asp Ser Leu Leu Ile Phe Lys Gly Val Asp Leu Val Arg Lys Thr Ser
805 810 815
Cys Asp Phe Val Lys Gly Val Val Lys Asp Ile Val Asp Leu Leu Phe
820 825 830
Phe Asp Glu Glu Val Gln Thr Ala Ala Val Glu Phe Ser His Met Thr
835 840 845
Gln Thr Gln Leu Arg Glu Gln Gly Val Pro Val Gly Ile His Lys Ile
850 855 860
Leu Arg Arg Leu Cys Glu Ala Arg Glu Glu Leu Phe Gln Asn Arg Ala
865 870 875 880
Asp Val Arg His Leu Met Leu Ser Ser Val Leu Ser Lys Glu Met Ala
885 890 895
Ala Tyr Lys Gln Pro Asn Leu Ala His Leu Ser Val Ile Arg Arg Leu
900 905 910
Ala Gln Arg Lys Glu Glu Ile Pro Asn Val Gly Asp Arg Ile Met Tyr
915 920 925
Val Leu Ile Ala Pro Ser Ile Gly Asn Lys Gln Thr His Asn Tyr Glu
930 935 940
Leu Ala Glu Asp Pro Asn Tyr Val Ile Glu His Lys Ile Pro Ile His
945 950 955 960

Ala Glu Lys Tyr Phe Asp Gln Ile Ile Lys Ala Val Thr Asn Ala Ile
965 970 975
Ser Pro Ile Phe Pro Lys Thr Asp Ile Lys Lys Glu Lys Leu Leu Leu
980 985 990
Tyr Leu Leu Pro Met Lys Val Tyr Leu Asp Glu Thr Phe Ser Ala Ile
995 1000 1005
Ala Glu Val Met
1010

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1013 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ser Ser Val Asn Leu Met Glu Trp Ser Ala Leu Lys Thr Gln Leu
1 5 10 15
Gln Ala Gly Arg Asp Ala Gly Lys Ala Arg Val Ser Ile Gly Pro Ala
20 25 30
Asp Thr Ala Arg Ile Thr Arg Met Thr Tyr Ala Asp Asn His Leu Ile
35 40 45
Val Phe Met Asn Ala Arg Leu Ala Lys Glu Asn His Arg Leu Tyr Gln
50 55 60
Phe Tyr Ala Glu Val Arg Cys Asp Leu Tyr Ser Tyr Lys Ser Cys Tyr
65 70 75 80
Gly Thr His Ala Ser Ala Thr Cys His Arg Asn Cys Thr Ser Tyr Lys
85 90 95
Thr Phe Val Met Pro Gly Leu Arg Asp Val His Thr Asp Lys Leu His
100 105 110
Val Val Lys Phe Lys Arg Ser Asp Glu Lys Arg Asp Lys Asn Cys Leu
115 120 125
Asp Gly Tyr Leu Ala Asp Val Asn Arg Val His Met Gln Thr Ser Leu
130 135 140
Leu Glu Gly Gln Tyr Val Arg Phe Lys Asn Ala His Ala Cys Arg Asp
145 150 155 160
Tyr Arg Leu Ser His Thr Ala Lys Asp Val His Glu Phe Glu Ser Met
165 170 175
Leu Glu Arg Val Gln Val Ser Ala Leu Ser His Glu Ile Leu Pro Val
180 185 190
Val Ala Cys Tyr Asp Ile Glu Thr His Ser Asp Gly Gln Arg Phe Ser
195 200 205
Ala Pro Asp Ala Asp Phe Ile Ile Ser Ile Ala Val Val Val Arg Arg
210 215 220
Asp Ala Ala Asp Thr Arg Ile Cys Leu Phe Tyr Ser Pro Asp Asp Pro
225 230 235 240
Val Asp Leu Ser Ser Ser Ser Ser Ser Pro Pro Ala Ala Pro Asp Thr
245 250 255
Ala Ala Val His Phe Arg Ala Glu Arg Asp Met Ile Ala Ala Phe Phe
260 265 270
Gln Leu Leu Pro Leu Leu Asn Ala Asp Val Val Leu Asp Phe Asn Gly
275 280 285
Asp Lys Phe Asp Leu Pro Phe Leu Thr Gly Arg Ala Asn Lys Leu Cys
290 295 300
Gly Pro Ala Glu Ala Ala Arg Ala Thr Lys Ile Ala Arg Tyr Asp Leu
305 310 315 320
Ser Pro Val Asn Val Val Thr Gln Gln Ser Tyr Asp Lys Phe Ser Asn
325 330 335
Lys Leu His Ser His Tyr Leu Thr Tyr Tyr Ile His Ile Asp Leu Tyr
340 345 350
Gln Phe Leu Ser Thr Asp Ser Glu His Asn Asp Leu Glu Asn Phe Gln

355 360 365
Leu Asn Thr Val Ala Glu His Tyr Leu Lys Lys Ser Lys Val Asp Leu
370 375 380
Pro Ile His Asp Met Leu Gln Met Tyr Gly Glu Lys Arg Leu Ser Arg
385 390 395 400
Ile Val Glu Tyr Asn Val Gln Asp Cys Val Leu Pro Val Glu Leu Phe
405 410 415
Leu Lys Leu Glu Ile Ala Asp Tyr Met Tyr Thr Gln Cys Met Leu Leu
420 425 430
Tyr Leu Cys Thr Asp Asp Leu Leu Arg Asn Ile Ser His Lys Ile Thr
435 440 445
Val Ala Tyr Phe His Leu Ala Leu Thr Asn Thr Val Ala Arg Arg Pro
450 455 460
Asp Pro Thr Pro Asp Pro Tyr Phe Phe Asn Lys Tyr Asp Leu Ser Val
465 470 475 480
Thr Ser Gly Ala Ser Ala Pro Ser Thr Ser Arg Pro Ala Asn Ala Ile
485 490 495
Asp Leu Ser Gln Leu Lys Arg Thr Pro Val Asp Ala Ala Arg Ile Pro
500 505 510
Pro Ser Ala Val Lys Leu Cys Ser Thr Arg Gln Ser Cys Thr Tyr Lys
515 520 525
Gly Gly Lys Val Leu Ser Pro Lys Pro Gly Phe Asn Arg Trp Val Ala
530 535 540
Thr Leu Asp Phe Asn Ala Leu Tyr Pro Thr Ile Met Met Trp Glu Gly
545 550 555 560
Val Cys Met Ser Ser Asn Val Phe Ile Ala Ser Asp Gly Asn Val Tyr
565 570 575
Leu Asp Lys Asn Val Asn Ala Val Asn Pro Lys Leu Leu Lys Thr Leu
580 585 590
Ser Glu Met Arg Val Arg Tyr Lys Gly Leu Arg Asp Gln Cys Glu Tyr
595 600 605
Asn Ser Phe Tyr Tyr Lys Leu Tyr Asp Lys Ile Gln Asn Ala Leu Lys
610 615 620
Arg Ile Ala Asn Ser Ile Tyr Gly Tyr Tyr Gly Ile Phe Phe Lys Pro
625 630 635 640
Leu Ala Asn Tyr Ile Thr Lys Met Gly Arg Gly Lys Leu Lys Glu Val
645 650 655
Val Gly Lys Val Glu Ala Met Ser Asp Asp Pro Arg Ile Leu Arg Glu
660 665 670
Phe Gly Leu Ser Lys Ile Asn Phe Ser Val Ile Tyr Gly Asp Thr Asp
675 680 685
Ser Cys Phe Ile Arg Val Leu Phe Asp Glu Ala Glu Trp Arg Arg Thr
690 695 700
Ala Ala Arg Pro Arg Ser Ala Pro Ser Cys Arg Thr Thr Cys Ala Lys
705 710 715 720
Arg Ser Thr Thr Leu Trp Cys Gly Tyr Lys Met Ser Leu Glu Asn Ile
725 730 735
Met Leu Ser Leu Ile Leu Leu Lys Lys Lys Lys Tyr Cys Tyr Leu Asn
740 745 750
Asn Glu Gln Arg Thr Lys Tyr Lys Gly Trp Leu Ile Lys Arg Asp Met
755 760 765
Pro Leu Phe Met Arg Lys Ala Phe Arg Ala Thr Val Asp Ser Phe Ser
770 775 780
Ala Ala Thr Arg Arg Val Arg Ala Arg Pro Ala Arg Arg Glu Met Leu
785 790 795 800
Arg Tyr Tyr Arg Glu Phe Gly Ala Pro Arg Glu Asn Leu Val Asp Tyr
805 810 815
Cys Phe Ser Met Ser Tyr Asn Glu Thr Ser Thr Thr Ala Lys Arg Arg
820 825 830
Lys Glu Glu Asp Pro Ala Arg Lys Pro Val Ile Thr Ile Ala Lys His
835 840 845
Cys Arg Glu Leu Leu Ala Asn Pro Gly Val Asp Phe Leu Pro Gly Asn
850 855 860
Gly Asp Arg Ile Gln Tyr Val Leu Val Asp Val Lys Glu Lys Ile Thr
865 870 875 880
Gln Lys Ala Phe Pro Leu Lys Leu Phe Asp Pro Asp Ser Pro Thr Leu
885 890 895
Gln Ile Ser Trp Leu Lys His Met Asn Ile Leu Cys Thr Phe Met Asn
900 905 910
Glu Leu Ile Gln Val Phe Gly Asn Arg Pro Glu Phe Glu His Tyr Phe

915 920 925
Gly Ala Ile Val Asp Glu Tyr Thr Ser Ala Gln Met Tyr Asp Val Arg
930 935 940
Tyr Pro Val Leu Val Pro Thr Arg Arg Ala Lys Ala Gly Lys Ser Ala
945 950 955 960
Lys Lys Asn Asp Ser Asp Ser Asp Ser Asp Ser Asp Asp Asp Asp Asp
965 970 975
Pro Ala Thr Thr Pro Val Asn Tyr His Ser Leu Phe Ser Met His Leu
980 985 990
Lys Lys Pro Lys Arg Gln Ala Val Gly Glu Phe Glu Pro Cys Pro Gln
995 1000 1005
Cys Val Ala Arg Ala
1010

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 985 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Asp Arg Asn Ala Val Leu Tyr Gly Val Leu Glu His Arg Leu Pro
1 5 10 15
Lys Trp Val Glu Leu Ser Asp Asp Thr Asp Leu Glu Pro Phe Phe Phe
20 25 30
Ser Ser Val Arg Tyr Ile Thr Ala Gly Ser Glu Asp Ala Ile Met Ile
35 40 45
Gln Ala Leu Asn Leu Asn Thr Asp Glu Ile Val Val Phe Leu Val Thr
50 55 60
Asn Leu Asn Phe Met Ala Leu Ile Pro Thr Val Tyr Ile Glu Asn Pro
65 70 75 80
Gly Ile Arg Gln Leu Ile Ala Ser Thr Pro Ile Ser Tyr Arg Ser Pro
85 90 95
Ile Thr Val Phe Asn Gly Asp Leu Lys Lys Trp Met Asp Cys Asp Leu
100 105 110
Phe Val Phe Gly Thr Met Ala Ala Gln Lys Ala Phe Ile Lys Ala Gly
115 120 125
Asn Ser Val Leu Gly Ser Leu Gly Gly Asn Val Tyr Thr Tyr Gly Asp
130 135 140
His Val Ser Asn Phe Asp Gly Asn Thr Pro Val Leu Gln Asn Asn Leu
145 150 155 160
Met Cys Ser His Val Tyr Tyr Thr Arg Tyr Lys Thr Asp Val Tyr Ala
165 170 175
Pro Trp Glu Phe Tyr Tyr Asp Gln Lys Arg Asp Gln Gly Tyr Leu Met
180 185 190
Ser Leu Pro Ala Ile Ile Pro Arg Cys Lys Arg Glu Gly Ala Phe Asp
195 200 205
Ile Glu Thr Ile Val His Glu Asn Ala Met Asp Gln Asp Leu Asn Cys
210 215 220
Gln Lys Phe Phe Lys Ser Glu Phe Arg Ser Met Glu Glu Ser Gln Val
225 230 235 240
Leu Ile Gln Arg Phe Arg Glu Ala Gly Val Thr Gly Leu Pro Pro Ser
245 250 255
Pro Phe Val Gly Ile Thr Gln Lys Leu His Glu Ile Val Ser Ile Ser
260 265 270
Leu Val Val Cys Asn Tyr His Lys Thr Gly Pro Lys Lys Lys Glu Tyr
275 280 285
Tyr Val Tyr Tyr Asn Thr Lys Lys Met Glu Asn Pro Met Glu Met Ile
290 295 300
Pro Val Glu His Leu His Leu Asp Ala Ser Arg Ile Lys Phe Glu Ala
305 310 315 320

Cys Lys Asn Glu Phe Tyr Met Leu Leu Ala Phe Ile Asn Arg Leu Arg
325 330 335
Lys Ser Val Asn Val Leu Tyr Val Tyr Asn Ala Gln Phe Asp Ile Gln
340 345 350
Val Ile Gln Gln Arg Leu Arg Tyr Tyr Ala Phe Lys Gln Arg Ala Pro
355 360 365
Arg Cys Cys Lys Gly His Asp Asp Ile Pro His Glu Trp Gly Lys Ala
370 375 380
Leu Met Glu Lys Trp Glu Ala Phe Leu Ser Val Lys Pro Gln Leu Phe
385 390 395 400
Lys Ala Gln Ile Leu Met Gly Gln Asp Ile Leu Lys Ala Asn Tyr Leu
405 410 415
Lys Leu Leu Glu Gly Ile Gly Ser Val Leu Ala Gln Ala Lys Ser Thr
420 425 430
Met Ala Lys Met Cys Thr Ile Lys Glu Arg Ile Asp Ser Tyr Arg Lys
435 440 445
Met Lys Asp Thr Val Gln Asn Phe Lys Ser His Gly Phe Gly Cys Asp
450 455 460
Ile Ile Asp Met Met Tyr Val Cys Lys Arg Lys Glu Phe Glu Ala Lys
465 470 475 480
Asp Gly Ser Leu Asn Thr Val Ala Gln Leu Ile Ile Lys Lys Phe Lys
485 490 495
Pro His Lys Ala Thr Pro Lys Ile His Lys Met Asp Asp Ile Thr Tyr
500 505 510
Asp Lys Leu Asp Gly Tyr Tyr Arg Ala Gly Gly Thr Lys Ile Ala Glu
515 520 525
Cys Leu Ile Tyr Asn Leu Ile Asp Ser Leu Leu Val Ile Arg Ile Ala
530 535 540
Lys Asn Leu Lys Pro Met Glu Glu Tyr Ile Tyr Arg Gln Leu Ala Cys
545 550 555 560
Tyr Asn Ile Asp Thr Ala Ala His Thr Arg Gly Val Met Asn Phe Cys
565 570 575
Gly Phe Ile Gln Ser Thr Lys Val Val Glu Val Ser Arg Asn Lys Ala
580 585 590
Arg Leu Asp Ala Gly Ile Val Met Ala Thr Asp Tyr Ile Arg Asn Ser
595 600 605
Leu Phe Thr Pro Glu Thr Ile Pro Arg Arg Gly Gly Phe Val Met Ala
610 615 620
Pro Leu Thr Gly Leu Phe Phe Ala Arg Pro Thr Gln Cys Phe Glu Leu
625 630 635 640
Cys Leu Asp Phe Thr Ser Met Tyr Pro Ser Met Met Cys Asp Leu Asn
645 650 655
Ile Ser Pro Glu Thr Ile Val Asp Ser Asp Lys Thr Asn Arg Val Gly
660 665 670
Asp Tyr Met Gly Tyr Asp Trp Ser Lys Ile Asp Gln Gly Phe Glu Lys
675 680 685
Phe Thr Leu Val Leu Arg Val Asp Arg Thr Asp Pro Glu Asn Pro Lys
690 695 700
Leu Val Arg His Thr Ser Asp Thr Ser Leu Ser Leu Lys Arg Tyr Leu
705 710 715 720
Arg Leu Arg Thr Glu His Lys Arg Ala Leu Lys Gln Ser Ser Gly Ser
725 730 735
Val Ala Glu Tyr His Asn Arg Leu Gln Asn Glu Met Lys Ile Cys Thr
740 745 750
Asn Thr His Tyr Gly Val Ser Glu His Thr Cys Ser Leu Met Ile Thr
755 760 765
Thr Gln Gly Gln His Lys Ile Lys Leu Val Asn Glu Phe Ile Lys Thr
770 775 780
Leu Asn Arg Thr Gly His Ser Leu Phe Pro Asn Tyr Gly Asp Thr Asp
785 790 795 800
Ser Thr Met Leu Tyr His Pro Ser Asp Glu Ser Glu Thr Gln Leu Glu
805 810 815
Asp Met Val Thr Leu Glu Asp Glu Met Arg Ala Glu Leu Arg Glu Tyr
820 825 830
Met Leu Lys Lys Leu Ser Ala Glu Leu Val Asn Arg Val Lys Glu Lys
835 840 845
Thr Lys Arg Thr Asp Thr Phe Val Gln Ser Phe Leu Ser Asp Val Glu
850 855 860
Thr Val Leu Phe Asp Asp Met Val Glu Lys Leu Arg Leu Phe Ser Gln
865 870 875 880

```
Gly Glu Val Ile Glu Pro Phe Lys Asp Gly Gly Thr Trp Trp Val Val
                885                 890                 895
Asp Pro Leu Thr Gly Ile Trp Met Asp Cys Ser Thr Pro Phe Ser Ser
            900                 905                 910
Glu Leu Ile Cys Lys Leu Glu Tyr Glu Asn Ala Ser Ser Ile Gly Cys
        915                 920                 925
His Val Ala Lys Lys Met Val Ser Ile Gly Ser Thr Tyr Leu Phe Phe
    930                 935                 940
Lys Lys Ile Ser Leu Tyr His Val Arg Val Trp Arg Met Cys Ala Asp
945                 950                 955                 960
Thr Asp Gly Ser Pro Ser His Leu Tyr Phe Pro Val Ser Leu Ser Arg
                965                 970                 975
Thr Arg Ala Lys Gln Arg Gly Asp His
            980                 985
```

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 964 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```
Met Lys Leu Lys Lys Leu Tyr Ile Phe Tyr Phe Asp Ile Tyr Glu Tyr
1               5                   10                  15
Phe Leu Cys Asp Leu Gln Leu Ser Glu Thr Asn Glu Ile Leu Lys Tyr
            20                  25                  30
Ile Lys Asn Asn Ile Asp Lys Tyr Thr Asn Ser Phe Asn Ser Ser Tyr
        35                  40                  45
Ile Ile Leu Lys Asp Phe Asn Ile Ile Thr Asn Glu Val Glu Leu Gln
    50                  55                  60
Ser Tyr Tyr Asn Phe Thr Glu Asp Ser Lys Ile Lys Leu Asn Asn Thr
65                  70                  75                  80
Asp Leu Ile Leu Phe Met Thr Pro Tyr Lys Ile Glu Arg Ile Tyr Ser
                85                  90                  95
Lys Tyr Asn Arg Asn Phe Asn Gln Tyr Arg Trp Phe Tyr Ile Leu Asn
            100                 105                 110
Asn Ile Glu Pro Ala Gly Ser Tyr Lys Ile Asn Met Ser Asn Leu Gln
        115                 120                 125
Asn Ile Asn Ile Tyr Asp Lys Asn Lys Thr Ala Tyr Tyr Cys Lys Asn
    130                 135                 140
Pro Lys Leu Leu Phe Leu Thr Pro Ile Glu Ile Asp Lys Phe Ile Pro
145                 150                 155                 160
Val Ser Arg Val Ser Ile Asp Ile Glu Cys Gln His Phe Gly Glu Phe
                165                 170                 175
Pro Thr Pro Asn Lys Phe Pro Ile Ser His Ile Cys Ile Asp Trp Phe
            180                 185                 190
Met Glu Ser Asn Ile Asn Pro Val Lys Lys Ile Ile Thr Leu Ile Asn
        195                 200                 205
Tyr Glu Ile Ile Lys Asn Tyr Lys Gly Glu Gln Lys Asp Arg Phe Ile
    210                 215                 220
Tyr Thr Glu Ile Asp Glu Leu Leu Thr Lys Asp Lys Val Tyr Ile Thr
225                 230                 235                 240
Ile Tyr Cys Thr Glu Lys Tyr Met Leu His Phe Ile Leu Tyr Thr Leu
                245                 250                 255
Arg Lys Asp Phe Asp Tyr Ile Leu Thr Tyr Asn Gly His Ser Phe Asp
            260                 265                 270
Phe Thr Tyr Ile Gln Gly Arg Arg Lys Phe Tyr Asn Leu Asn Glu Leu
        275                 280                 285
Cys Leu Val Asn Ala His Lys Ser Asn Glu Leu Lys Ile Tyr Ser Tyr
    290                 295                 300
Asn Lys Asp Thr Thr Tyr Glu Ile Asp Ser Asn Asn Gly Ile Ile Phe
```

```
305                 310                 315                 320
Leu Asp Leu Tyr Asn Tyr Ile Lys Lys Ile Tyr Asn Tyr Asn Ser Tyr
                325                 330                 335
Lys Leu Gly Glu Ile Ala Lys Glu Arg Phe Asn Ile Leu Ser Lys Ile
            340                 345                 350
Ile Asp Asn Gly Asp Glu Tyr Ile Ile Met Pro Leu Asp Thr Ala Asp
        355                 360                 365
Asn Lys Asn Lys Val Ser Ile Phe Tyr Asp Val Ile Arg Thr Ala Asn
    370                 375                 380
Tyr Cys Phe Ile Asn Asn Ile Pro Tyr Lys Ile Lys Asp Lys Thr Lys
385                 390                 395                 400
Ile Ile Asn Asp Lys Glu Lys Leu Tyr Asp Pro Ile Ser Ile Glu Asn
                405                 410                 415
Ser Leu Tyr Gln Gln Phe Lys Ile Tyr Lys Asn Asn Thr Pro Ile Ser
            420                 425                 430
Asp Glu Thr Thr Lys Val Met Leu Ser Lys Asp Asp Val Asp Ile Gly
        435                 440                 445
Asn Lys Asn Ala Tyr Val Asn Phe Thr Lys Asp Lys Ser Asp Asp Ile
    450                 455                 460
Ala Tyr Tyr Cys Thr His Asp Thr Val Leu Cys Asn Cys Ile Phe Lys
465                 470                 475                 480
Tyr Asp Met Ile His Asp Lys Val Ile Ala Phe Ser Asn Glu Tyr Leu
                485                 490                 495
Leu Pro Gln Tyr Met Ser Phe Lys Tyr Lys Ser Thr Thr Asn Ile Ser
            500                 505                 510
Gly Leu Leu Leu Lys Thr Leu Phe Cys Asn Arg Ser Met Ile Val Ser
        515                 520                 525
Gly Asn Leu Glu Phe Glu Lys Phe Glu Gly Gly Tyr Val Leu Glu Pro
    530                 535                 540
Lys Gln Lys Tyr Ile Asp Ser Ile Thr Ala Val Phe Asp Phe Asn Ser
545                 550                 555                 560
Glu Tyr Pro Ser Asn Ile Ile Glu Ala Asn Leu Ser Pro Glu Lys Val
                565                 570                 575
Glu Lys Val Ile Lys Leu Gln Asp Asp Glu Tyr Ala Val Asp Ile Val
            580                 585                 590
Glu Asn Tyr Leu Lys Glu Lys Tyr Pro Tyr Pro Asp Tyr Cys Tyr Met
        595                 600                 605
Leu Ile Lys Lys Asp Lys Thr Tyr Lys Phe Ile Val Met Asp Arg Arg
    610                 615                 620
Lys Pro Gly Ile Ile Thr Gln Met Ile Asp Lys Gly Met Lys Ser Lys
625                 630                 635                 640
Asn Glu Tyr Lys Asn Leu Lys Asn Ile Asn Lys Asn Asn Pro Val Leu
                645                 650                 655
Tyr Asn Tyr Tyr Thr Ser Ala Leu Tyr Ser Lys Lys Ile Thr Ile Asn
            660                 665                 670
Ser Leu Tyr Gly Leu Leu Gly Ser Glu Arg Phe Asp Phe Asn Ser Pro
        675                 680                 685
Tyr Cys Ala Glu Tyr Cys Thr Ala Leu Gly Gln Lys Cys Ile Lys Tyr
    690                 695                 700
Ile Lys Asn Leu Val Asp Lys Ser Arg Tyr Ile Asp Asn Asn Leu Tyr
705                 710                 715                 720
Leu Asn Glu Gln Asn Asn Pro Phe Ser Asn Glu Pro Val Ile Thr Arg
                725                 730                 735
Tyr Ser Gly Asn Leu Asp Val Asn Phe Thr Phe Tyr Ile Ile Tyr Gly
            740                 745                 750
Asp Thr Asp Ser Leu Phe Ile Asn Ile Lys Phe Asp Asn Lys Phe Asp
        755                 760                 765
Asn Lys Glu Asp Leu Val Asn Lys Ser His Glu Cys Phe Gln Phe Leu
    770                 775                 780
Ser Asn Ile Ile Asn Asp Glu Lys Asn Ile Ile Leu Ser Lys Asn Phe
785                 790                 795                 800
Asn Phe Glu Tyr Glu Lys Met Tyr Ile Trp Met Leu Leu Leu Ala Lys
                805                 810                 815
Lys Lys Tyr Ile Gly Glu Val Val Ser Ser Met Asn Pro Leu Gln Leu
            820                 825                 830
Ile Ser Asp Ser Lys Gly Thr Ala Leu Ile Arg Arg Asp Cys Thr Glu
        835                 840                 845
Ile His Lys Thr Ile Leu Lys Asn Thr Ile Asp Ile Leu Lys Glu Tyr
    850                 855                 860
Leu Thr Asn Asn Cys Thr Ile Gln Asp Val Asn Asn Lys Ile Asn Asn
```

```
865                 870                 875                 880
Tyr Leu Met Phe Thr Phe Lys Asn Ile Ile Glu Asn Ile Gln Asn Leu
                885                 890                 895
Asp Ile Asn Glu Phe Lys Lys Ser Val Lys Tyr Thr Gly Ile Tyr Lys
            900                 905                 910
Asp Pro Asn Phe Tyr Ile Glu Leu Cys Val Lys Lys Tyr Asn Ser Lys
        915                 920                 925
Asn Pro Asn Asp Lys Ile Val Lys Gly Gln Arg Phe Asp Phe Ile Tyr
    930                 935                 940
Ala His Glu Ile Asp Ile Trp Asp Ile Glu Thr Lys Lys Trp Asn Thr
945                 950                 955                 960
Lys Tyr Thr Ser
```

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 763 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```
Met Pro Leu Ser Tyr Gln His Phe Arg Lys Leu Leu Leu Leu Asp Asp
 1               5                  10                  15
Glu Thr Glu Ala Gly Pro Leu Glu Glu Glu Leu Pro Arg Leu Ala Asp
            20                  25                  30
Ala Asp Leu Asn Arg Arg Val Ala Glu Asp Leu Asn Leu Gly Asn Leu
        35                  40                  45
Asn Val Ser Ile Pro Trp Thr His Lys Val Gly Asn Phe Thr Gly Leu
    50                  55                  60
Tyr Ser Ser Thr Val Pro Ile Phe Asn Pro Glu Trp Gln Thr Pro Ser
65                  70                  75                  80
Phe Pro Lys Ile His Leu His Glu Asp Ile Ala Asn Arg Cys Gln Gln
                85                  90                  95
Phe Val Gly Pro Leu Thr Val Asn Glu Lys Arg Arg Leu Lys Leu Ile
            100                 105                 110
Met Pro Ala Arg Phe Tyr Pro Asn Ser Thr Lys Tyr Leu Pro Leu Asp
        115                 120                 125
Lys Gly Ile Lys Thr Tyr Tyr Pro Asp His Val Val Asn His Tyr Phe
    130                 135                 140
Gln Thr Arg His Tyr Leu His Thr Leu Trp Lys Ala Gly Ile Leu Tyr
145                 150                 155                 160
Lys Arg Glu Thr Thr Arg Ser Ala Ser Phe Cys Gly Ser Pro Tyr Ser
                165                 170                 175
Trp Glu Gln Glu Leu His His Gly Arg Leu Val Ile Lys Thr Ser Gln
            180                 185                 190
Arg His Gly Asp Glu Pro Phe Cys Ser Gln Pro Ser Gly Ile Leu Ser
        195                 200                 205
Arg Ser Ser Val Gly Pro Cys Ile Arg Ser Gln Phe Lys Gln Ser Arg
    210                 215                 220
Leu Gly Leu Gln Pro His Gln Gly Pro Leu Ala Thr Ser Gln Pro Gly
225                 230                 235                 240
Arg Ser Gly Ser Ile Trp Ala Arg Val His Ser Pro Thr Arg Arg Cys
                245                 250                 255
Phe Gly Val Glu Pro Ser Gly Ser Gly His Il  Gly His Arg Ala Ser
            260                 265                 270
Asp Ala Ser Ser Cys Leu His Gln Ser Ala Val Arg Lys Ala Ala Tyr
        275                 280                 285
Ser His Leu Ser Thr Ser Lys Arg Gln Ser Ser Ser Gly His Ala Val
    290                 295                 300
Glu Phe His Ser Phe Pro Pro Ser Ser Ala Arg Ser Gln Ser Gln Gly
305                 310                 315                 320
```

```
Pro Val Phe Ser Cys Trp Trp Leu Gln Phe Arg Asn Thr Gln Pro Cys
                325                 330                 335
Ser Asn Tyr Cys Leu Ser His Leu Val Asn Leu Leu Glu Asp Trp Gly
            340                 345                 350
Pro Cys Thr Glu His Gly Glu His His Ile Arg Ile Pro Arg Thr Pro
        355                 360                 365
Ala Arg Val Thr Gly Gly Val Phe Leu Val Asp Lys Asn Pro His Asn
    370                 375                 380
Thr Ala Glu Ser Arg Leu Val Val Asp Phe Ser Gln Phe Ser Arg Gly
385                 390                 395                 400
Ser Thr Arg Val Ser Trp Pro Lys Phe Ala Val Pro Asn Leu Gln Ser
                405                 410                 415
Leu Thr Asn Leu Leu Ser Ser Asn Leu Ser Trp Leu Ser Leu Asp Val
            420                 425                 430
Ser Ala Ala Phe Tyr His Ile Pro Leu His Pro Ala Ala Met Pro His
        435                 440                 445
Leu Leu Ile Gly Ser Ser Gly Leu Ser Arg Tyr Val Ala Arg Leu Ser
    450                 455                 460
Ser Asn Ser Arg Ile Asn Asn Asn Gln His Gly Thr Leu Gln Asn Leu
465                 470                 475                 480
His Asp Ser Cys Ser Arg Gln Leu Tyr Val Ser Leu Met Leu Leu Tyr
                485                 490                 495
Lys Thr Tyr Gly Trp Lys Leu His Leu Tyr Ser His Pro Ile Ile Leu
            500                 505                 510
Gly Phe Arg Lys Ile Pro Met Gly Val Gly Leu Ser Pro Phe Leu Leu
        515                 520                 525
Ala Gln Phe Thr Ser Ala Ile Cys Ser Val Val Arg Arg Ala Phe Pro
    530                 535                 540
His Cys Leu Ala Phe Ser Tyr Met Asp Asp Val Val Leu Gly Ala Lys
545                 550                 555                 560
Ser Val Gln His Leu Glu Ser Leu Tyr Thr Ala Val Thr Asn Phe Leu
                565                 570                 575
Leu Ser Leu Gly Ile His Leu Asn Pro Asn Lys Thr Lys Arg Trp Gly
            580                 585                 590
Tyr Ser Leu Asn Phe Met Gly Tyr Val Ile Gly Ser Trp Gly Thr Leu
        595                 600                 605
Pro Gln Asp His Ile Val Gln Lys Ile Lys His Cys Phe Arg Lys Leu
    610                 615                 620
Pro Val Asn Arg Pro Ile Asp Trp Lys Val Cys Gln Arg Leu Val Gly
625                 630                 635                 640
Leu Leu Gly Phe Ala Ala Pro Phe Thr Gln Cys Gly Tyr Pro Ala Leu
                645                 650                 655
Met Pro Leu Tyr Ala Cys Ile Gln Ala Lys Gln Ala Phe Thr Phe Ser
            660                 665                 670
Pro Thr Tyr Lys Ala Phe Leu Ser Lys Gln Tyr Met Asn Leu Tyr Pro
        675                 680                 685
Val Ala Arg Gln Arg Pro Gly Leu Cys Gln Val Phe Ala Asp Ala Thr
    690                 695                 700
Pro Thr Gly Trp Gly Leu Ala Ile Gly His Gln Arg Met Arg Glu Thr
705                 710                 715                 720
Phe Val Ala Pro Leu Pro Ile His Thr Ala Glu Leu Leu Ala Ala Cys
                725                 730                 735
Phe Ala Arg Ser Arg Ser Gly Ala Lys Leu Ile Gly Thr Asp Asn Ser
            740                 745                 750
Val Val Leu Ser Gln Lys Tyr Thr Ser Phe Pro
        755                 760
```

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1663 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

```
Met Gly Pro Thr Ser Gly Ser Gln Leu Leu Val Leu Leu Leu Leu Leu
 1               5                  10                  15
Ala Ser Ser Leu Leu Ala Leu Gly Ser Pro Met Tyr Ser Ile Ile Thr
            20                  25                  30
Pro Asn Val Leu Arg Leu Glu Ser Glu Glu Thr Phe Ile Leu Glu Ala
        35                  40                  45
His Asp Ala Gln Gly Asp Val Pro Val Thr Val Thr Val Gln Asp Phe
    50                  55                  60
Leu Lys Lys Gln Val Leu Thr Ser Glu Lys Thr Val Leu Thr Gly Ala
65                  70                  75                  80
Thr Gly His Leu Asn Arg Val Phe Ile Lys Ile Pro Ala Ser Lys Glu
                85                  90                  95
Phe Asn Ala Asp Lys Gly His Lys Tyr Val Thr Val Val Ala Asn Phe
            100                 105                 110
Gly Ala Thr Val Val Glu Lys Ala Val Leu Val Ser Phe Gln Ser Gly
        115                 120                 125
Tyr Leu Phe Ile Gln Thr Asp Lys Thr Ile Tyr Thr Pro Gly Ser Thr
    130                 135                 140
Val Phe Tyr Arg Ile Phe Thr Val Asp Asn Asn Leu Leu Pro Val Gly
145                 150                 155                 160
Lys Thr Val Val Ile Val Ile Glu Thr Pro Asp Gly Val Pro Ile Lys
                165                 170                 175
Arg Asp Ile Leu Ser Ser His Asn Gln Tyr Gly Ile Leu Pro Leu Ser
            180                 185                 190
Trp Asn Ile Pro Glu Leu Val Asn Met Gly Gln Trp Lys Ile Arg Ala
        195                 200                 205
Phe Tyr Glu His Ala Pro Lys Gln Thr Phe Ser Ala Glu Phe Glu Val
    210                 215                 220
Lys Glu Tyr Val Leu Pro Ser Phe Glu Val Leu Val Glu Pro Thr Glu
225                 230                 235                 240
Lys Phe Tyr Tyr Ile His Gly Pro Lys Gly Leu Glu Val Ser Ile Thr
                245                 250                 255
Ala Arg Phe Leu Tyr Gly Lys Asn Val Asp Gly Thr Ala Phe Val Ile
            260                 265                 270
Phe Gly Val Gln Asp Glu Asp Lys Lys Ile Ser Leu Ala Leu Ser Leu
        275                 280                 285
Thr Arg Val Leu Ile Glu Asp Gly Ser Gly Glu Ala Val Leu Ser Arg
    290                 295                 300
Lys Val Leu Met Asp Gly Val Arg Pro Ser Ser Pro Glu Ala Leu Val
305                 310                 315                 320
Gly Lys Ser Leu Tyr Val Ser Val Thr Val Ile Leu His Ser Gly Ser
                325                 330                 335
Asp Met Val Glu Ala Glu Arg Ser Gly Ile Pro Ile Val Thr Ser Pro
            340                 345                 350
Tyr Gln Ile His Phe Thr Lys Thr Pro Lys Phe Phe Lys Pro Ala Met
        355                 360                 365
Pro Phe Asp Leu Met Val Phe Val Thr Asn Pro Asp Gly Ser Pro Ala
    370                 375                 380
Arg Arg Val Pro Val Val Thr Gln Gly Ser Asp Ala Gln Ala Leu Thr
385                 390                 395                 400
Gln Asp Asp Gly Val Ala Lys Leu Ser Val Asn Thr Pro Asn Asn Arg
                405                 410                 415
Gln Pro Leu Thr Ile Thr Val Ser Thr Lys Lys Glu Gly Ile Pro Asp
            420                 425                 430
Ala Arg Gln Ala Thr Arg Thr Met Gln Ala Gln Pro Tyr Ser Thr Met
        435                 440                 445
His Asn Ser Asn Asn Tyr Leu His Leu Ser Val Ser Arg Val Glu Leu
    450                 455                 460
Lys Pro Gly Asp Asn Leu Asn Val Asn Phe His Leu Arg Thr Asp Ala
465                 470                 475                 480
Gly Gln Glu Ala Lys Ile Arg Tyr Tyr Thr Tyr Leu Val Met Asn Lys
                485                 490                 495
Gly Lys Leu Leu Lys Ala Gly Arg Gln Val Arg Glu Pro Gly Gln Asp
            500                 505                 510
Leu Val Val Leu Ser Leu Pro Ile Thr Pro Glu Phe Ile Pro Ser Phe
        515                 520                 525
Arg Leu Val Ala Tyr Tyr Thr Leu Ile Gly Ala Asn Gly Gln Arg Glu
    530                 535                 540
```

Val Val Ala Asp Ser Val Trp Val Asp Val Lys Asp Ser Cys Val Gly
545 550 555 560
Thr Leu Val Val Lys Gly Asp Pro Arg Asp Asn Arg Gln Pro Ala Pro
565 570 575
Gly His Gln Thr Thr Leu Arg Ile Glu Gly Asn Gln Gly Ala Arg Val
580 585 590
Gly Leu Val Ala Val Asp Lys Gly Val Phe Val Leu Asn Lys Lys Asn
595 600 605
Lys Leu Thr Gln Ser Lys Ile Trp Asp Val Val Glu Lys Ala Asp Ile
610 615 620
Gly Cys Thr Pro Gly Ser Gly Lys Asn Tyr Ala Gly Val Phe Met Asp
625 630 635 640
Ala Gly Leu Thr Phe Lys Thr Asn Gln Gly Leu Gln Thr Asp Gln Arg
645 650 655
Glu Asp Pro Glu Cys Ala Lys Pro Ala Ala Arg Arg Arg Arg Ser Val
660 665 670
Gln Leu Met Glu Arg Arg Met Asp Lys Ala Gly Gln Tyr Thr Asp Lys
675 680 685
Gly Leu Arg Lys Cys Cys Glu Asp Gly Met Arg Asp Ile Pro Met Pro
690 695 700
Tyr Ser Cys Gln Arg Arg Ala Arg Leu Ile Thr Gln Gly Glu Ser Cys
705 710 715 720
Leu Lys Ala Phe Met Asp Cys Cys Asn Tyr Ile Thr Lys Leu Arg Glu
725 730 735
Gln His Arg Arg Asp His Val Leu Gly Leu Ala Arg Ser Asp Val Asp
740 745 750
Glu Asp Ile Ile Pro Glu Glu Asp Ile Ile Ser Arg Ser His Phe Pro
755 760 765
Glu Ser Trp Leu Trp Thr Ile Glu Glu Leu Lys Glu Pro Glu Lys Asn
770 775 780
Gly Ile Ser Thr Lys Val Met Asn Ile Phe Leu Lys Asp Ser Ile Thr
785 790 795 800
Thr Trp Glu Ile Leu Ala Val Ser Leu Ser Asp Lys Lys Gly Ile Cys
805 810 815
Val Ala Asp Pro Tyr Glu Ile Thr Val Met Gln Asp Phe Phe Ile Asp
820 825 830
Leu Arg Leu Pro Tyr Ser Val Val Arg Asn Glu Gln Val Glu Ile Arg
835 840 845
Ala Val Leu Phe Asn Tyr Arg Glu Gln Glu Lys Leu Lys Val Arg Val
850 855 860
Glu Leu Leu His Asn Pro Ala Phe Cys Ser Met Ala Thr Ala Lys Lys
865 870 875 880
Arg Tyr Tyr Gln Thr Ile Glu Ile Pro Pro Lys Ser Ser Val Ala Val
885 890 895
Pro Tyr Val Ile Val Pro Leu Lys Ile Gly Leu Gln Glu Val Glu Val
900 905 910
Lys Ala Ala Val Phe Asn His Phe Ile Ser Asp Gly Val Lys Lys Ile
915 920 925
Leu Lys Val Val Pro Glu Gly Met Arg Val Asn Lys Thr Val Ala Val
930 935 940
Arg Thr Leu Asp Pro Glu His Leu Asn Gln Gly Gly Val Gln Arg Glu
945 950 955 960
Asp Val Asn Ala Ala Asp Leu Ser Asp Gln Val Pro Asp Thr Asp Ser
965 970 975
Glu Thr Arg Ile Leu Leu Gln Gly Thr Pro Val Ala Gln Met Ala Glu
980 985 990
Asp Ala Val Asp Gly Glu Arg Leu Lys His Leu Ile Val Thr Pro Ser
995 1000 1005
Gly Cys Gly Glu Gln Asn Met Ile Gly Met Thr Pro Thr Val Ile Ala
1010 1015 1020
Val His Tyr Leu Asp Gln Thr Glu Gln Trp Glu Lys Phe Gly Leu Glu
025 1030 1035 1040
Lys Arg Gln Glu Ala Leu Glu Leu Ile Lys Lys Gly Tyr Thr Gln Gln
1045 1050 1055
Leu Ala Phe Lys Gln Pro Ile Ser Ala Tyr Al Ala Phe Asn Asn Arg
1060 1065 1070
Pro Pro Ser Thr Trp Leu Thr Ala Met Trp Ser Arg Ser Phe Ser Leu
1075 1080 1085
Ala Ala Asn Leu Ile Ala Ile Asp Ser Gln Val Leu Cys Gly Ala Val
1090 1095 1100

Lys Trp Leu Ile Leu Glu Lys Gln Lys Pro Asp Gly Val Phe Gln Glu
105 1110 1115 1120
Asp Gly Pro Val Ile His Gln Glu Met Ile Gly Gly Phe Arg Asn Thr
1125 1130 1135
Lys Glu Ala Asp Val Ser Leu Thr Ala Phe Val Leu Ile Ala Leu Gln
1140 1145 1150
Glu Ala Arg Asp Ile Cys Glu Gly Gln Val Asn Ser Leu Pro Gly Ser
1155 1160 1165
Ile Asn Lys Ala Gly Glu Tyr Leu Glu Ala Ser Tyr Leu Asn Leu Gln
1170 1175 1180
Arg Pro Tyr Thr Val Ala Ile Ala Gly Tyr Ala Leu Ala Leu Met Asn
185 1190 1195 1200
Lys Leu Glu Glu Pro Tyr Leu Thr Lys Phe Leu Asn Thr Ala Lys Asp
1205 1210 1215
Arg Asn Arg Trp Glu Glu Pro Gly Gln Gln Leu Tyr Asn Val Glu Ala
1220 1225 1230
Thr Ser Tyr Ala Leu Leu Ala Leu Leu Leu Leu Lys Asp Phe Asp Ser
1235 1240 1245
Val Pro Pro Val Val Arg Trp Leu Asn Asp Glu Arg Tyr Tyr Gly Gly
1250 1255 1260
Gly Tyr Gly Ser Thr Gln Ala Thr Phe Met Val Phe Gln Ala Leu Ala
265 1270 1275 1280
Gln Tyr Arg Ala Asp Val Pro Asp His Lys Asp Leu Asn Met Asp Val
1285 1290 1295
Ser Leu His Leu Pro Ser Arg Ser Ser Pro Thr Val Phe Arg Leu Leu
1300 1305 1310
Trp Glu Ser Gly Ser Leu Leu Arg Ser Glu Glu Thr Lys Gln Asn Glu
1315 1320 1325
Gly Phe Ser Leu Thr Ala Lys Gly Lys Gly Gln Gly Thr Leu Ser Val
1330 1335 1340
Val Thr Val Tyr His Ala Lys Val Lys Gly Lys Thr Thr Cys Lys Lys
345 1350 1355 1360
Phe Asp Leu Arg Val Thr Ile Lys Pro Ala Pro Glu Thr Ala Lys Lys
1365 1370 1375
Pro Gln Asp Ala Lys Ser Ser Met Ile Leu Asp Ile Cys Thr Arg Tyr
1380 1385 1390
Leu Gly Asp Val Asp Ala Thr Met Ser Ile Leu Asp Ile Ser Met Met
1395 1400 1405
Thr Gly Phe Ile Pro Asp Thr Asn Asp Leu Glu Leu Leu Ser Ser Gly
1410 1415 1420
Val Asp Arg Tyr Ile Ser Lys Tyr Glu Met Asp Lys Ala Phe Ser Asn
425 1430 1435 1440
Lys Asn Thr Leu Ile Ile Tyr Leu Glu Lys Ile Ser His Ser Glu Glu
1445 1450 1455
Asp Cys Leu Ser Phe Lys Val His Gln Phe Phe Asn Val Gly Leu Ile
1460 1465 1470
Gln Pro Gly Ser Val Lys Val Tyr Ser Tyr Tyr Asn Leu Glu Glu Ser
1475 1480 1485
Cys Thr Arg Phe Tyr His Pro Glu Lys Asp Asp Gly Met Leu Ser Lys
1490 1495 1500
Leu Cys His Asn Glu Met Cys Arg Cys Ala Glu Glu Asn Cys Phe Met
505 1510 1515 1520
His Gln Ser Gln Asp Gln Val Ser Leu Asn Glu Arg Leu Asp Lys Ala
1525 1530 1535
Cys Glu Pro Gly Val Asp Tyr Val Tyr Lys Thr Lys Leu Thr Thr Ile
1540 1545 1550
Glu Leu Ser Asp Asp Phe Asp Glu Tyr Ile Met Thr Ile Glu Gln Val
1555 1560 1565
Ile Lys Ser Gly Ser Asp Glu Val Gln Ala Gly Gln Glu Arg Arg Phe
1570 1575 1580
Ile Ser His Val Lys Cys Arg Asn Ala Leu Lys Leu Gln Lys Gly Lys
585 1590 1595 1600
Gln Tyr Leu Met Trp Gly Leu Ser Ser Asp Leu Trp Gly Glu Lys Pro
1605 1610 1615
Asn Thr Ser Tyr Ile Ile Gly Lys Asp Thr Trp Val Glu His Trp Pro
1620 1625 1630
Glu Ala Glu Glu Arg Gln Asp Gln Lys Asn Gln Lys Gln Cys Glu Asp
1635 1640 1645
Leu Gly Ala Phe Thr Glu Thr Met Val Val Phe Gly Cys Pro Asn
1650 1655 1660

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1666 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

```
Met Gly Pro Ala Ala Gly Pro Ser Leu Leu Leu Leu Leu Leu Ala Ser
 1               5                  10                  15
Val Ser Leu Ala Leu Gly Asp Pro Met Tyr Ser Ile Ile Thr Pro Asn
            20                  25                  30
Ile Leu Arg Leu Glu Asn Glu Glu Thr Val Val Leu Glu Ala His Glu
        35                  40                  45
Val Gln Gly Asp Ile Pro Val Thr Val Thr Val His Asp Phe Pro Ala
    50                  55                  60
Lys Lys Asn Val Leu Ser Ser Glu Lys Thr Val Leu Thr Ser Ala Thr
65                  70                  75                  80
Gly Tyr Leu Gly Thr Val Thr Ile Lys Ile Pro Ala Ser Lys Glu Phe
                85                  90                  95
Lys Ser Asp Lys Gly Arg Lys Leu Val Val Val Gln Ala Ala Phe Gly
            100                 105                 110
Gly Thr Gln Leu Glu Lys Val Val Leu Val Ser Leu Gln Ser Gly Tyr
        115                 120                 125
Leu Phe Ile Gln Thr Asp Lys Thr Ile Tyr Thr Pro Gly Ser Thr Val
    130                 135                 140
Leu Tyr Arg Ile Phe Thr Val Asp Ser Asp Leu Leu Pro Val Gly Arg
145                 150                 155                 160
Thr Ile Ile Val Thr Ile Glu Thr Pro Asp Gly Ile Pro Ile Lys Arg
                165                 170                 175
Asp Thr Leu Ser Ser Asn Asn Gln His Gly Ile Leu Pro Leu Ser Trp
            180                 185                 190
Asn Ile Pro Glu Leu Val Asn Met Gly Gln Trp Lys Ile Gln Ala Phe
        195                 200                 205
Tyr Glu Asn Ser Pro Lys Gln Val Phe Ser Ala Glu Phe Glu Val Lys
    210                 215                 220
Glu Tyr Val Leu Pro Ser Phe Glu Val Leu Val Glu Pro Thr Glu Lys
225                 230                 235                 240
Phe Tyr Tyr Ile Asp Asp Pro Lys Gly Leu Glu Val Asn Ile Ile Ala
                245                 250                 255
Arg Phe Leu Tyr Gly Lys Asn Val Asp Gly Thr Ala Phe Val Ile Phe
            260                 265                 270
Gly Val Gln Asp Gly Asp Gln Arg Ile Ser Leu Ala Gln Ser Leu Thr
        275                 280                 285
Arg Val Val Ile Glu Asp Gly Ser Gly Glu Val Val Leu Ser Arg Gln
    290                 295                 300
Val Leu Leu Asp Gly Val Gln Pro Ser Arg Pro Glu Ala Leu Val Gly
305                 310                 315                 320
Lys Ser Leu Tyr Val Ser Val Thr Val Ile Leu His Ser Gly Ser Asp
                325                 330                 335
Met Val Glu Ala Glu Arg Ser Gly Ile Pro Ile Val Thr Ser Pro Tyr
            340                 345                 350
Gln Ile His Phe Thr Lys Thr Pro Lys Tyr Phe Lys Pro Ala Met Pro
        355                 360                 365
Phe Glu Ile Met Val Leu Val Thr Asn Pro Asp Gly Ser Pro Ala Pro
    370                 375                 380
His Val Pro Val Val Thr Gln Gly Ser Asn Val Gln Ser Leu Thr Gln
385                 390                 395                 400
Ala Asp Gly Val Ala Arg Leu Ser Ile Asn Thr Pro Asn Thr Arg Gln
                405                 410                 415
Pro Leu Ser Val Thr Val Gln Thr Lys Lys Gly Gly Ile Pro Asp Ala
            420                 425                 430
```

Arg Gln Ala Ile Asn Thr Met Gln Ala Leu Pro Tyr Thr Thr Met Tyr
435 440 445
Asn Ser Asn Asn Tyr Leu His Leu Ser Met Pro Arg Thr Glu Leu Lys
450 455 460
Pro Gly Glu Thr Ile Asn Val Asn Phe His Leu Arg Ser Asp Pro Asn
465 470 475 480
Gln Glu Ala Lys Ile Arg Tyr Tyr Thr Tyr Leu Ile Met Asn Lys Gly
485 490 495
Lys Leu Leu Lys Val Gly Arg Gln Pro Arg Glu Pro Gly Gln Ala Leu
500 505 510
Val Val Leu Pro Met Pro Ile Thr Lys Glu Leu Ile Pro Ser Phe Arg
515 520 525
Leu Val Ala Tyr Tyr Thr Leu Ile Gly Ala Ser Ala Gln Arg Glu Val
530 535 540
Val Ala Asp Ser Val Trp Ala Asp Val Arg Asp Ser Cys Val Gly Thr
545 550 555 560
Leu Val Val Lys Gly Gly Ser Gly Lys Asp Gly Gln Asp Lys Arg Gln
565 570 575
Gln His Leu Pro Arg Gln Gln Met Thr Leu Arg Ile Glu Gly Asn Gln
580 585 590
Gly Ala Arg Val Gly Leu Val Ala Val Asp Lys Gly Val Phe Val Leu
595 600 605
Asn Lys Lys His Lys Leu Thr Gln Ser Lys Ile Trp Asp Val Val Glu
610 615 620
Lys Ala Asp Ile Gly Cys Thr Pro Gly Ser Gly Lys Asp Tyr Ala Gly
625 630 635 640
Val Phe Thr Asp Ala Gly Leu Ser Phe Lys Ser Ser Lys Ala Gly Leu
645 650 655
Gln Thr Ala Gln Arg Glu Gly Leu Asp Cys Pro Lys Pro Ala Ala Arg
660 665 670
Arg Arg Arg Ser Val Gln Leu Met Glu Arg Arg Met Asp Lys Ala Gly
675 680 685
Lys Tyr Lys Ser Lys Glu Leu Arg Arg Cys Cys Glu Asp Gly Met Arg
690 695 700
Glu Asn Pro Met Gln Phe Ser Cys Gln Arg Arg Ala Arg Tyr Val Ser
705 710 715 720
Leu Gly Glu Ala Cys Val Lys Ala Phe Leu Asp Cys Cys Thr Tyr Met
725 730 735
Ala Gln Leu Arg Gln Gln His Arg Arg Glu Gln Asn Leu Gly Leu Ala
740 745 750
Arg Ser Asp Met Asp Glu Asp Ile Ile Pro Glu Glu Asp Ile Ile Ser
755 760 765
Arg Ser Gln Phe Pro Glu Ser Trp Leu Trp Thr Ile Glu Glu Leu Lys
770 775 780
Glu Pro Glu Arg Asn Gly Ile Ser Thr Lys Thr Met Asn Ile Phe Leu
785 790 795 800
Lys Asp Ser Ile Thr Thr Trp Glu Ile Leu Ala Val Ser Leu Ser Asp
805 810 815
Lys Lys Gly Ile Cys Val Ala Asp Pro Phe Glu Val Thr Val Met Gln
820 825 830
Asp Phe Phe Ile Asp Leu Arg Leu Pro Tyr Ser Val Val Arg Asn Glu
835 840 845
Gln Val Glu Ile Arg Ala Val Leu Tyr Asn Tyr Arg Glu Ala Gln Ser
850 855 860
Leu Lys Val Arg Val Glu Leu Leu His Asn Pro Ala Phe Cys Ser Leu
865 870 875 880
Ala Thr Ala Lys Lys Arg His Thr Gln Thr Val Thr Ile Gly Pro Lys
885 890 895
Ser Ser Val Ala Val Pro Tyr Val Leu Val Pro Leu Lys Ile Gly Leu
900 905 910
Gln Glu Val Glu Val Lys Ala Ala Val Tyr Asn Tyr Phe Ile Ser Asp
915 920 925
Gly Val Lys Lys Thr Leu Lys Val Val Pro Glu Gly Met Arg Val Asn
930 935 940
Lys Thr Val Ala Ile Arg Thr Leu Asn Pro Glu Gln Leu Gly Gln Gly
945 950 955 960
Gly Val Gln Arg Glu Glu Ile Pro Ala Ala Asp Leu Ser Asp Gln Val
965 970 975
Pro Asp Thr Asp Ser Glu Thr Lys Ile Leu Leu Gln Gly Thr Pro Val
980 985 990

```
Ala Gln Met Ala Glu Asp Ala Val Asp Ala Glu Arg Leu Lys His Leu
        995                1000                1005
Ile Ile Thr Pro Ser Gly Cys Gly Glu Gln Asn Met Ile Gly Met Thr
   1010                1015                1020
Pro Thr Val Ile Ala Val His Tyr Leu Asp Gln Thr Glu Gln Trp Glu
025                1030                1035                1040
Lys Phe Gly Leu Glu Lys Arg Gln Glu Ala Leu Asn Leu Ile Asn Arg
               1045                1050                1055
Gly Tyr Thr Gln Gln Leu Ala Phe Lys Gln Pro Asn Trp Ala Tyr Ala
           1060                1065                1070
Ala Phe Lys Asn Arg Ala Ser Ser Thr Trp Leu Thr Ala Tyr Val Val
       1075                1080                1085
Lys Val Phe Ser Leu Ala Ala Asn Leu Ile Gly Ile Asp Ser Glu Val
   1090                1095                1100
Leu Cys Gly Ala Val Lys Trp Leu Ile Leu Glu Lys Gln Lys Pro Asp
105                1110                1115                1120
Gly Val Phe Gln Glu Asp Gly Pro Val Ile His Gln Glu Met Ile Gly
               1125                1130                1135
Gly Val Arg Thr Ala Gln Glu Ala Asp Val Ser Leu Thr Ala Phe Val
           1140                1145                1150
Leu Ile Ala Leu Gln Glu Ala Lys Asp Ile Cys Arg Ala Gln Val Asn
       1155                1160                1165
Asn Leu Glu Ala Asn Ile Asn Lys Ala Gly Asp Tyr Ile Glu Ser Arg
   1170                1175                1180
Tyr Ala Asp Val Arg Arg Pro Tyr Thr Leu Ala Ile Ala Gly Tyr Ala
185                1190                1195                1200
Leu Ala Leu Leu Glu Arg Leu Asn Gly Ala Thr Leu Gln Lys Phe Leu
               1205                1210                1215
Asn Ala Ala Thr Glu Lys Asn Arg Trp Glu Glu Ala Arg Gln Lys Leu
           1220                1225                1230
Tyr Ser Val Glu Ala Thr Ser Tyr Ala Leu Leu Ala Leu Leu Leu Leu
       1235                1240                1245
Lys Asp Phe Asp Ala Val Pro Pro Val Val Arg Trp Leu Asn Glu Gln
   1250                1255                1260
Arg Tyr Tyr Gly Arg Gly Tyr Gly Ser Thr Gln Ala Thr Phe Met Val
265                1270                1275                1280
Phe Gln Ala Leu Ala Gln Tyr Gln Thr Asp Val Pro Asp His Lys Asp
               1285                1290                1295
Leu Asn Met Glu Val Ala Leu Gln Leu Pro Ser Arg Ser Ser Pro Ser
           1300                1305                1310
Lys Phe Arg Leu Val Trp Glu Ala Gly Ser Leu Leu Arg Ser Glu Ala
       1315                1320                1325
Thr Lys Gln Asn Glu Gly Phe Lys Leu Thr Ala Lys Gly Lys Gly Gln
   1330                1335                1340
Gly Thr Leu Ser Val Val Ala Val Tyr Tyr Ala Lys Thr Lys Arg Lys
345                1350                1355                1360
Val Val Cys Lys Asn Phe Asp Leu Arg Val Thr Leu Lys Pro Ala Pro
               1365                1370                1375
Asp Thr Val Lys Lys Pro Gln Glu Ala Lys Ser Thr Met Ile Leu Gly
           1380                1385                1390
Ile Cys Thr Arg Tyr Leu Gly Asp Gln Asp Ala Thr Met Ser Ile Leu
       1395                1400                1405
Asp Ile Ser Met Met Thr Gly Phe Ile Pro Asp Thr Asp Asp Leu Lys
   1410                1415                1420
Leu Leu Ala Thr Gly Val Asp Arg Tyr Ile Ser Lys Tyr Glu Met Asn
425                1430                1435                1440
Lys Asp Phe Ser Lys Asn Thr Leu Ile Ile Tyr Leu Asp Lys Val Ser
               1445                1450                1455
His Ser Glu Glu Glu Cys Leu Ser Phe Lys Ile His Gln Phe Phe Asn
           1460                1465                1470
Val Gly Leu Ile Gln Pro Gly Ser Val Lys Val Tyr Ser Tyr Tyr Asn
       1475                1480                1485
Leu Asp Glu Thr Cys Thr Gln Phe Tyr His Pro Glu Lys Glu Asp Gly
   1490                1495                1500
Met Leu Asn Lys Leu Cys His Lys Asp Leu Cys Arg Cys Ala Glu Glu
505                1510                1515                1520
Asn Cys Phe Ile Gln Leu Pro Glu Lys Ile Thr Leu Asp Glu Arg Leu
               1525                1530                1535
Glu Lys Ala Cys Glu Pro Gly Val Asp Tyr Val Tyr Lys Thr Lys Leu
           1540                1545                1550
```

```
Leu Lys Met Glu Leu Ser Asp Asp Phe Asp Glu Tyr Ile Met Thr Ile
        1555                1560                1565
Glu Gln Val Ile Lys Ser Gly Ser Asp Glu Val Gln Ala Gly Lys Glu
    1570                1575                1580
Arg Arg Phe Ile Ser His Ile Lys Cys Arg Asp Ala Leu His Leu Lys
585                 1590                1595                1600
Glu Gly Lys His Tyr Leu Met Trp Gly Leu Ser Ser Asp Leu Trp Gly
                1605                1610                1615
Glu Arg Pro Asn Met Ser Tyr Ile Ile Gly Lys Asp Thr Trp Val Glu
            1620                1625                1630
Ala Trp Pro Glu Ala Glu Glu Cys Gln Asp Glu Glu Asn Gln Gln Gln
        1635                1640                1645
Cys Gln Asp Leu Gly Thr Phe Thr Glu Asn Met Val Val Phe Gly Cys
    1650                1655                1660
Pro Asn
665
```

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1015 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

```
Phe Phe Arg Glu Asp Leu Ala Phe Leu Gln Gly Lys Ala Arg Glu Phe
1               5                   10                  15
Ser Ser Glu Gln Thr Arg Ala Asn Ser Pro Thr Ile Ser Ser Glu Gln
            20                  25                  30
Thr Arg Ala Asn Ser Pro Thr Arg Arg Glu Leu Gln Val Trp Gly Arg
        35                  40                  45
Asp Asn Asn Ser Pro Ser Glu Ala Gly Ala Asp Arg Gln Gly Thr Val
    50                  55                  60
Ser Phe Asn Phe Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Val Thr
65                  70                  75                  80
Ile Lys Ile Gly Gly Gln Leu Lys Glu Ala Leu Leu Asp Thr Gly Ala
                85                  90                  95
Asp Asp Thr Val Leu Glu Glu Met Ser Leu Pro Gly Arg Trp Lys Pro
            100                 105                 110
Lys Met Ile Gly Gly Ile Gly Gly Phe Ile Lys Val Arg Gln Tyr Asp
        115                 120                 125
Gln Ile Leu Ile Glu Ile Cys Gly His Lys Ala Ile Gly Thr Val Leu
    130                 135                 140
Val Gly Pro Thr Pro Val Asn Ile Ile Gly Arg Asn Leu Leu Thr Gln
145                 150                 155                 160
Ile Gly Cys Thr Leu Asn Phe Pro Ile Ser Pro Ile Glu Thr Val Pro
                165                 170                 175
Val Lys Leu Lys Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro
            180                 185                 190
Leu Thr Glu Glu Lys Ile Lys Ala Leu Val Glu Ile Cys Thr Glu Met
        195                 200                 205
Glu Lys Glu Gly Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn
    210                 215                 220
Thr Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys
225                 230                 235                 240
Leu Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu
                245                 250                 255
Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser
            260                 265                 270
Val Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp
        275                 280                 285
Glu Asp Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn
```

290 295 300
Glu Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp
305 310 315 320
Lys Gly Ser Pro Ala Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu
325 330 335
Pro Phe Lys Lys Gln Asn Pro Asp Ile Val Ile Tyr Gln Tyr Met Asp
340 345 350
Asp Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys
355 360 365
Ile Glu Glu Leu Arg Gln His Leu Leu Arg Trp Gly Leu Thr Thr Pro
370 375 380
Asp Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu
385 390 395 400
Leu His Pro Asp Lys Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys
405 410 415
Asp Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn
420 425 430
Trp Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys
435 440 445
Leu Leu Arg Gly Thr Lys Ala Leu Thr Glu Val Ile Pro Leu Thr Glu
450 455 460
Glu Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro
465 470 475 480
Val His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile
485 490 495
Gln Lys Gln Gly Gln Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro
500 505 510
Phe Lys Asn Leu Lys Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His
515 520 525
Thr Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Thr Thr
530 535 540
Glu Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile
545 550 555 560
Gln Lys Glu Thr Trp Glu Thr Trp Trp Thr Glu Tyr Trp Gln Ala Thr
565 570 575
Trp Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu
580 585 590
Trp Tyr Gln Leu Glu Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr
595 600 605
Val Asp Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr
610 615 620
Val Thr Asn Lys Gly Arg Gln Lys Val Val Pro Leu Thr Asn Thr Thr
625 630 635 640
Asn Gln Lys Thr Glu Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser
645 650 655
Gly Leu Glu Val Asn Ile Val Thr Asp Ser Gln Tyr Ala Leu Gly Ile
660 665 670
Ile Gln Ala Gln Pro Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile
675 680 685
Ile Glu Gln Leu Ile Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro
690 695 700
Ala His Lys Gly Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser
705 710 715 720
Ala Gly Ile Arg Lys Ile Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln
725 730 735
Asp Glu His Glu Lys Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp
740 745 750
Phe Asn Leu Pro Pro Val Val Ala Lys Glu Ile Val Ala Ser Cys Asp
755 760 765
Lys Cys Gln Leu Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser
770 775 780
Pro Gly Ile Trp Gln Leu Asp Cys Thr His Leu Glu Gly Lys Val Ile
785 790 795 800
Leu Val Ala Val His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile
805 810 815
Pro Ala Glu Thr Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala
820 825 830
Gly Arg Trp Pro Val Lys Thr Ile His Thr Asp Asn Gly Ser Asn Phe
835 840 845
Thr Ser Ala Thr Val Lys Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln

```
    850                 855                 860
Glu Phe Gly Ile Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser
865                 870                 875                 880
Met Asn Lys Glu Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala
                885                 890                 895
Glu His Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe
            900                 905                 910
Lys Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Val
        915                 920                 925
Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile
    930                 935                 940
Thr Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asn Pro
945                 950                 955                 960
Leu Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val
                965                 970                 975
Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala
            980                 985                 990
Lys Ile Ile Arg Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val
        995                 1000                1005
Ala Ser Arg Gln Asp Glu Asp
    1010                1015
```

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1034 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

```
Thr Gly Gly Phe Phe Arg Asp Trp Pro Leu Gly Lys Glu Ala Pro Gln
1               5                   10                  15
Phe Pro Arg Gly Pro Ser Ser Thr Gly Ala Asn Thr Asn Ser Thr Pro
            20                  25                  30
Ile Gly Ser Ser Ser Gly Ser Thr Gly Glu Ile Tyr Ala Ala Arg Glu
        35                  40                  45
Lys Ala Glu Gly Ala Glu Thr Glu Thr Ile Gln Arg Gly Asp Arg Gly
    50                  55                  60
Leu Thr Ala Pro Arg Thr Arg Arg Gly Pro Met Gln Gly Asp Asn Arg
65                  70                  75                  80
Gly Leu Ala Ala Pro Gln Phe Ser Leu Trp Lys Arg Pro Val Val Thr
                85                  90                  95
Ala His Ile Glu Gly Gln Pro Val Glu Val Leu Leu Asp Thr Gly Ala
            100                 105                 110
Asp Asp Ser Ile Val Ala Gly Ile Glu Leu Gly Ser Asn Tyr Ser Pro
        115                 120                 125
Lys Ile Val Gly Gly Ile Gly Gly Phe Ile Asn Thr Lys Glu Tyr Lys
    130                 135                 140
Asn Val Glu Ile Glu Val Leu Gly Lys Arg Val Arg Ala Thr Ile Met
145                 150                 155                 160
Thr Gly Asp Thr Pro Ile Asn Ile Phe Gly Arg Asn Ile Leu Thr Ala
                165                 170                 175
Leu Gly Met Ser Leu Asn Leu Pro Val Ala Lys Ile Glu Pro Ile Lys
            180                 185                 190
Ile Met Leu Lys Pro Gly Lys Asp Gly Pro Arg Leu Arg Gln Trp Pro
        195                 200                 205
Leu Thr Lys Glu Lys Ile Glu Ala Leu Lys G[illegible] Ile Cys Glu Lys Met
    210                 215                 220
Glu Lys Glu Gly Gln Leu Glu Glu Ala Pro Pro Thr Asn Pro Tyr Asn
225                 230                 235                 240
Thr Pro Thr Phe Ala Ile Arg Lys Lys Asp Lys Asn Lys Trp Arg Met
                245                 250                 255
```

Leu Ile Asp Phe Arg Glu Leu Asn Lys Val Thr Gln Asp Phe Thr Glu
260 265 270
Ile Gln Leu Gly Ile Pro His Pro Ala Gly Leu Ala Lys Lys Arg Arg
275 280 285
Ile Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Ile Pro Leu His
290 295 300
Glu Asp Phe Arg Gln Tyr Thr Ala Phe Thr Leu Pro Ser Val Asn Asn
305 310 315 320
Ala Glu Pro Gly Lys Arg Tyr Ile Tyr Lys Val Leu Pro Gln Gly Trp
325 330 335
Lys Gly Ser Pro Ala Ile Phe Gln Tyr Thr Met Arg Gln Val Leu Glu
340 345 350
Pro Phe Arg Lys Ala Asn Ser Asp Val Ile Ile Ile Gln Tyr Met Asp
355 360 365
Asp Ile Leu Ile Ala Ser Asp Arg Thr Asp Leu Glu His Asp Lys Val
370 375 380
Val Leu Gln Leu Lys Glu Leu Leu Asn Asn Leu Gly Phe Ser Thr Pro
385 390 395 400
Asp Glu Lys Phe Gln Lys Asp Pro Pro Tyr Arg Trp Met Gly Tyr Glu
405 410 415
Leu Trp Pro Thr Lys Trp Lys Leu Gln Lys Ile Gln Leu Pro Gln Lys
420 425 430
Glu Val Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Val Leu Asn
435 440 445
Trp Ala Ala Gln Ile Tyr Pro Gly Ile Lys Thr Lys His Leu Cys Arg
450 455 460
Leu Ile Arg Gly Lys Met Thr Leu Thr Glu Glu Val Gln Trp Thr Glu
465 470 475 480
Leu Ala Glu Ala Glu Leu Glu Glu Asn Arg Ile Ile Leu Ser Gln Glu
485 490 495
Gln Glu Gly His Tyr Tyr Gln Glu Glu Lys Glu Leu Glu Ala Thr Val
500 505 510
Gln Lys Asp Gln Asp Asn Gln Trp Thr Tyr Lys Ile His Gln Glu Glu
515 520 525
Lys Ile Leu Lys Val Gly Lys Tyr Ala Lys Ile Lys His Thr His Thr
530 535 540
Asn Gly Val Lys Leu Leu Ala Gln Val Val Gln Lys Ile Gly Lys Glu
545 550 555 560
Ala Leu Val Ile Gly Arg Ile Pro Lys Phe His Leu Pro Val Glu Arg
565 570 575
Glu Val Trp Glu Gln Trp Trp Asp Asn Tyr Trp Gln Val Thr Trp Ile
580 585 590
Pro Asp Trp Asp Phe Val Ser Thr Pro Pro Leu Val Arg Leu Ala Phe
595 600 605
Asn Leu Val Gly Asp Pro Ile Pro Gly Thr Glu Thr Phe Tyr Thr Asp
610 615 620
Gly Ser Cys Asn Arg Gln Ser Lys Glu Gly Lys Ala Gly Tyr Val Thr
625 630 635 640
Asp Arg Gly Arg Asp Lys Val Lys Ile Leu Glu Gln Thr Thr Asn Gln
645 650 655
Gln Ala Glu Leu Glu Ala Phe Ala Met Ala Leu Thr Asp Ser Gly Pro
660 665 670
Lys Ala Asn Ile Ile Val Asp Ser Gln Tyr Val Met Gly Ile Val Ala
675 680 685
Gly Gln Pro Thr Glu Ser Glu Asn Arg Ile Val Asn Gln Ile Ile Glu
690 695 700
Glu Met Ile Lys Lys Glu Ala Ile Tyr Val Ala Trp Val Pro Ala His
705 710 715 720
Lys Gly Ile Gly Gly Asn Gln Glu Val Asp His Leu Val Ser Gln Gly
725 730 735
Ile Arg Gln Val Leu Phe Leu Glu Lys Ile Glu Pro Ala Gln Glu Glu
740 745 750
His Glu Lys Tyr His Thr Asn Val Lys Glu Leu Cys His Lys Phe Asp
755 760 765
Ile Pr Gln Leu Val Ala Arg Gln Ile Val Asn Thr Cys Ala Gln Tyr
770 775 780
Gln Gln Lys Gly Glu Ala Ile His Gly Gln Val Asn Ala Glu Val Gly
785 790 795 800
Thr Trp Gln Met Asp Cys Thr His Leu Glu Gly Lys Ile Ile Ile Val
805 810 815

Ala Val His Val Ala Ser Gly Phe Ile Glu Ala Glu Val Ile Pro Gln
820 825 830
Glu Ser Gly Arg Gln Thr Ala Leu Phe Leu Leu Lys Leu Ala Ser Arg
835 840 845
Trp Pro Ile Thr His Leu His Thr Asp Asn Gly Ala Asn Phe Thr Ser
850 855 860
Gln Glu Val Lys Met Val Ala Trp Trp Val Gly Ile Glu Gln Thr Phe
865 870 875 880
Gly Val Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ala Met Asn
885 890 895
His His Leu Lys Asn Gln Ile Ser Arg Ile Arg Glu Gln Ala Asn Thr
900 905 910
Val Glu Thr Ile Val Leu Met Ala Val His Cys Met Asn Phe Lys Arg
915 920 925
Arg Gly Gly Ile Gly Asp Met Thr Pro Ser Glu Arg Leu Ile Asn Met
930 935 940
Ile Thr Thr Glu Gln Glu Ile Gln Phe Leu Gln Ala Lys Asn Ser Lys
945 950 955 960
Leu Lys Asn Phe Arg Val Tyr Phe Arg Glu Gly Arg Asp Gln Leu Trp
965 970 975
Lys Gly Pro Gly Glu Leu Leu Trp Lys Gly Asp Gly Ala Val Ile Val
980 985 990
Lys Val Gly Thr Asp Ile Lys Ile Ile Pro Arg Arg Lys Ala Lys Ile
995 1000 1005
Ile Arg Asp Tyr Gly Gly Arg Gln Glu Leu Asp Ser Ser Ser His Leu
1010 1015 1020
Glu Gly Ala Arg Glu Asn Gly Glu Val Ala
025 1030 1

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1022 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Pro Arg Lys Thr Ser Gly Phe Phe Arg Ala Trp Pro Met Gly Lys
1 5 10 15
Glu Ala Pro Gln Phe Pro His Gly Pro Asp Ala Ser Gly Ala Asp Thr
20 25 30
Asn Cys Ser Pro Arg Gly Ser Ser Cys Gly Ser Thr Glu Glu Leu His
35 40 45
Glu Asp Gly Gln Lys Ala Glu Gly Glu Gln Arg Glu Thr Leu Gln Gly
50 55 60
Gly Asn Gly Gly Phe Ala Ala Pro Gln Phe Ser Leu Trp Arg Arg Pro
65 70 75 80
Ile Val Thr Ala Tyr Ile Glu Glu Gln Pro Val Glu Val Leu Leu Asp
85 90 95
Thr Gly Ala Asp Asp Ser Ile Val Ala Gly Ile Glu Leu Gly Pro Asn
100 105 110
Tyr Thr Pro Lys Ile Val Gly Gly Ile Gly Gly Phe Ile Asn Thr Lys
115 120 125
Glu Tyr Lys Asp Val Lys Ile Lys Val Leu Gly Lys Val Ile Lys Gly
130 135 140
Thr Ile Met Thr Gly Asp Thr Pro Ile Asn Ile Phe Gly Arg Asn Leu
145 150 155 160
Leu Thr Ala Met Gly Met Ser Leu Asn Leu Pro Ile Ala Lys Val Glu
165 170 175
Pro Ile Lys Val Thr Leu Lys Pro Gly Lys Asp Gly Pro Lys Leu Arg
180 185 190
Gln Trp Pro Leu Ser Lys Glu Lys Ile Ile Ala Leu Arg Glu Ile Cys

195 200 205
Glu Lys Met Glu Lys Asp Gly Gln Leu Glu Glu Ala Pro Pro Thr Asn
210 215 220
Pro Tyr Asn Thr Pro Thr Phe Ala Ile Lys Lys Lys Asp Lys Asn Lys
225 230 235 240
Trp Arg Met Leu Ile Asp Phe Arg Glu Leu Asn Lys Val Thr Gln Asp
245 250 255
Phe Thr Glu Val Gln Leu Gly Ile Pro His Pro Ala Gly Leu Ala Lys
260 265 270
Arg Arg Arg Ile Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Ile
275 280 285
Pro Leu Asp Glu Glu Phe Arg Gln Tyr Thr Ala Phe Thr Leu Pro Ser
290 295 300
Val Asn Asn Ala Glu Pro Gly Lys Arg Tyr Ile Tyr Lys Val Leu Pro
305 310 315 320
Gln Gly Trp Lys Gly Ser Pro Ala Ile Phe Gln His Thr Met Arg Asn
325 330 335
Val Leu Glu Pro Phe Arg Lys Ala Asn Pro Asp Val Thr Leu Ile Gln
340 345 350
Tyr Met Asp Asp Ile Leu Ile Ala Ser Asp Arg Thr Asp Leu Glu His
355 360 365
Asp Arg Val Val Leu Gln Leu Lys Glu Leu Leu Asn Ser Ile Gly Phe
370 375 380
Ser Thr Pro Glu Glu Lys Phe Gln Lys Asp Pro Pro Phe Gln Trp Met
385 390 395 400
Gly Tyr Glu Leu Trp Pro Thr Lys Trp Lys Leu Gln Lys Ile Glu Leu
405 410 415
Pro Gln Arg Glu Thr Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly
420 425 430
Val Leu Asn Trp Ala Ala Gln Ile Tyr Pro Gly Ile Lys Thr Lys His
435 440 445
Leu Cys Arg Leu Ile Arg Gly Lys Met Thr Leu Thr Glu Glu Val Gln
450 455 460
Trp Thr Glu Met Ala Glu Ala Glu Tyr Glu Glu Asn Lys Ile Ile Leu
465 470 475 480
Ser Gln Glu Gln Glu Gly Cys Tyr Tyr Gln Glu Gly Lys Pro Leu Glu
485 490 495
Ala Thr Val Ile Lys Ser Gln Asp Asn Gln Trp Ser Tyr Lys Ile His
500 505 510
Gln Glu Asp Lys Ile Leu Lys Val Gly Lys Phe Ala Lys Ile Lys Asn
515 520 525
Thr His Thr Asn Gly Val Arg Leu Leu Ala His Val Val Gln Lys Ile
530 535 540
Gly Lys Glu Ala Ile Val Ile Trp Gly Gln Val Pro Arg Phe His Leu
545 550 555 560
Pro Val Glu Arg Glu Ile Trp Glu Gln Trp Trp Thr Asp Tyr Trp Gln
565 570 575
Val Thr Trp Ile Pro Glu Trp Asp Phe Val Ser Thr Pro Pro Leu Val
580 585 590
Arg Leu Val Phe Asn Leu Val Lys Glu Pro Ile Gln Gly Ala Glu Thr
595 600 605
Phe Tyr Val Asp Gly Ser Cys Asn Arg Gln Ser Arg Glu Gly Lys Ala
610 615 620
Gly Tyr Val Thr Asp Arg Gly Arg Asp Lys Ala Lys Leu Leu Glu Gln
625 630 635 640
Thr Thr Asn Gln Gln Ala Glu Leu Glu Ala Phe Tyr Leu Ala Leu Ala
645 650 655
Asp Ser Gly Pro Lys Ala Asn Ile Ile Val Asp Ser Gln Tyr Val Met
660 665 670
Gly Ile Val Ala Gly Gln Pro Thr Glu Ser Glu Ser Arg Leu Val Asn
675 680 685
Gln Ile Ile Glu Glu Met Ile Lys Lys Glu Ala Ile Tyr Val Ala Trp
690 695 700
Val Pro Ala His Lys Gly Ile Gly Gly Asn Gl[illegible] Glu Val Asp His Leu
705 710 7[illegible] 720
Val Ser Gln Gly Ile Arg Gln Val Leu Phe Le[illegible] Glu Lys Ile Glu Pro
725 730 735
Ala Gln Glu Glu His Glu Lys Tyr His Ser Asn Val Lys Glu Leu Val
740 745 750
Phe Lys Phe Gly Leu Pro Arg Leu Val Ala Lys Gln Ile Val Asp Thr

```
        755                     760                     765
Cys Asp Lys Cys His Gln Lys Gly Glu Ala Ile His Gly Gln Val Asn
    770                     775                     780
Ala Glu Leu Gly Thr Trp Gln Met Asp Cys Thr His Leu Glu Gly Lys
785                     790                     795                 800
Ile Ile Ile Val Ala Val His Val Ala Ser Gly Phe Ile Glu Ala Glu
                805                     810                     815
Val Ile Pro Gln Glu Thr Gly Arg Gln Thr Ala Leu Phe Leu Leu Lys
            820                     825                     830
Leu Ala Ser Arg Trp Pro Ile Thr His Leu His Thr Asp Asn Gly Ala
        835                     840                     845
Asn Phe Thr Ser Gln Glu Val Lys Met Val Ala Trp Trp Ala Gly Ile
    850                     855                     860
Glu Gln Thr Phe Gly Val Pro Tyr Asn Pro Gln Ser Gln Gly Val Val
865                     870                     875                 880
Glu Ala Met Asn His His Leu Lys Thr Gln Ile Asp Arg Ile Arg Glu
                885                     890                     895
Gln Ala Asn Ser Ile Glu Thr Ile Val Leu Met Ala Val His Cys Met
            900                     905                     910
Asn Phe Lys Arg Arg Gly Gly Ile Gly Asp Met Thr Pro Ala Glu Arg
        915                     920                     925
Leu Val Asn Met Ile Thr Thr Glu Gln Glu Ile Gln Phe Gln Gln Ser
    930                     935                     940
Lys Asn Ser Lys Phe Lys Asn Phe Arg Val Tyr Tyr Arg Glu Gly Arg
945                     950                     955                 960
Asp Gln Leu Trp Lys Gly Pro Gly Glu Leu Leu Trp Lys Gly Glu Gly
                965                     970                     975
Ala Val Ile Leu Lys Val Gly Thr Glu Ile Lys Val Val Pro Arg Arg
            980                     985                     990
Lys Ala Lys Ile Ile Lys Asp Tyr Gly Gly Gly Lys Glu Leu Asp Ser
        995                    1000                    1005
Gly Ser His Leu Glu Asp Thr Gly Glu Ala Arg Glu Val Ala
   1010                    1015                    1020
```

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1027 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
Ser Thr Lys Lys Lys Arg Leu Leu Ala Val Trp Ala Arg Gly Thr Pro
1               5                   10                  15
Asn Glu Arg Leu His Arg Lys Thr Gly Glu Phe Phe Arg Glu Arg Leu
            20                  25                  30
Ala Phe Pro Gln Arg Glu Ala Arg Gln Leu Cys Ala Glu Gln Asn Arg
        35                  40                  45
Thr Asn Gly Pro Thr Asp Arg Glu Leu Trp Val Pro Gly Gly Arg Glu
    50                  55                  60
Glu Pro Gly Glu Glu Arg Gly Arg Glu Gln Ser Ile Ser Thr Asn Leu
65                  70                  75                  80
Pro Gln Ile Thr Leu Trp Gln Arg Pro Leu Ile Pro Val Lys Val Glu
                85                  90                  95
Gly Gln Leu Cys Glu Ala Leu Leu Asp Thr Gly Ala Asp Asp Thr Val
            100                 105                 110
Ile Glu Arg Ile Gln Leu Gln Gly Leu Trp Lys Pro Lys Met Ile Gly
        115                 120                 125
Gly Ile Gly Gly Phe Ile Lys Val Lys Gln Phe Asp Asn Val His Ile
    130                 135                 140
Glu Ile Glu Gly Arg Lys Val Val Gly Thr Val Leu Val Gly Pro Thr
145                 150                 155                 160
```

Pro Val Asn Ile Ile Gly Arg Asn Ile Leu Thr Gln Leu Gly Cys Thr
165 170 175
Leu Val Phe Pro Ile Ser Ser Ile Glu Thr Val Pro Val Lys Leu Lys
180 185 190
Pro Gly Met Asp Gly Pro Lys Val Lys Gln Trp Pro Leu Ser Ala Glu
195 200 205
Lys Ile Lys Ala Leu Thr Glu Ile Cys Gln Glu Met Glu Lys Glu Gly
210 215 220
Lys Ile Ser Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Ile Phe
225 230 235 240
Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe
245 250 255
Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly
260 265 270
Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu
275 280 285
Asp Val Gly Asp Ala Tyr Phe Ser Cys Pro Leu Asp Lys Asp Phe Arg
290 295 300
Lys Tyr Thr Ala Phe Thr Ile Pro Ser Ile Asn Asn Glu Thr Pro Gly
305 310 315 320
Val Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro
325 330 335
Ser Ile Phe Gln Ser Ser Met Thr Lys Ile Leu Glu Pro Phe Arg Glu
340 345 350
Lys Asn Pro Asp Ile Thr Ile Tyr Gln Tyr Met Asp Asp Leu Tyr Val
355 360 365
Gly Ser Asp Leu Glu Ile Asp Gln His Arg Lys Lys Val Glu Glu Leu
370 375 380
Arg Gln His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp Lys Lys His
385 390 395 400
Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp
405 410 415
Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Glu Val Trp Thr
420 425 430
Val Asn Asp Ile Gln Lys Leu Ile Gly Lys Leu Asn Trp Ala Ser Gln
435 440 445
Ile Tyr Pro Gly Ile Lys Ile Lys Gln Leu Cys Lys Leu Ile Arg Gly
450 455 460
Thr Lys Lys Leu Thr Asp Val Val Pro Leu Thr Pro Glu Ala Glu Leu
465 470 475 480
Glu Leu Ala Glu Asn Arg Glu Ile Val Ser Thr Pro Val His Gly Val
485 490 495
Tyr Tyr Asp Pro Asp Lys Glu Leu Ile Ala Glu Ile Gln Lys Gln Gly
500 505 510
Asn Cys Gln Trp Thr Tyr Gln Ile Phe Gln Glu Pro His Lys Asn Leu
515 520 525
Lys Thr Gly Lys Tyr Ala Arg Gln Arg Ser Ala His Thr Asn Asp Ile
530 535 540
Arg Gln Leu Ala Glu Ala Val Gln Lys Ile Ala Thr Glu Ser Ile Val
545 550 555 560
Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Val Gln Lys Glu Ser
565 570 575
Trp Glu Ala Trp Trp Ala Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu
580 585 590
Trp Glu Phe Ile Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Ser Leu
595 600 605
Glu Thr Glu Pro Ile Pro Thr Thr Asp Thr Tyr Tyr Val Asp Gly Ala
610 615 620
Ala Asn Arg Glu Thr Lys Thr Gly Lys Ala Gly Tyr Val Thr Asp Lys
625 630 635 640
Gly Lys Gln Lys Ile Ile Ser Leu Glu Asn Thr Thr Asn Gln Gln Ala
645 650 655
Glu Leu Lys Ala Leu Leu Leu Ala Leu Gln Asp Ser Asp Gln Gln Val
660 665 670
Asn Ile Val Thr Asp Ser Gln Tyr Val Leu Gly Ile Ile Gln Ser Gln
675 680 685
Pro Asp His Ser Glu S r Glu Leu Val Asn Gln Ile Ile Glu Glu Leu
690 695 700
Ile Lys Lys Glu Lys Ile Tyr Leu Ser Trp Val Pro Ala His Lys Gly
705 710 715 720

Ile Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ala Gly Ile Arg
725 730 735
Lys Val Leu Phe Leu Asp Gly Ile Asp Arg Ala Gln Glu Glu His Glu
740 745 750
Arg Tyr His Ser Asn Trp Lys Ala Met Ala Ser Asp Phe Asn Leu Pro
755 760 765
Pro Ile Val Ala Lys Glu Ile Val Ala His Cys Asp Lys Cys Gln Val
770 775 780
Lys Gly Glu Ala Met His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp
785 790 795 800
Gln Val Asp Cys Thr His Leu Glu Gly Lys Val Ile Ile Val Ala Val
805 810 815
His Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr
820 825 830
Gly Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro
835 840 845
Val Lys Thr Ile His Thr Asp Asn Gly Pro Asn Phe Thr Ser Ala Ala
850 855 860
Val Lys Ala Ala Cys Trp Trp Ala Asp Ile Lys Gln Glu Phe Gly Ile
865 870 875 880
Pro Tyr Asn Pro Gln Ser Gln Gly Val Val Glu Ser Leu Asn Lys Glu
885 890 895
Leu Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys
900 905 910
Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly
915 920 925
Gly Ile Gly Gly Tyr Thr Ala Gly Glu Arg Ile Ile Asp Ile Ile Ala
930 935 940
Thr Asp Ile Gln Thr Ser Glu Leu Gln Lys Gln Ile Leu Lys Val Gln
945 950 955 960
Lys Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Ile Trp Lys Gly
965 970 975
Pro Ala Thr Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp
980 985 990
Gln Gly Glu Leu Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg
995 1000 1005
Asp Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln
1010 1015 1020
Asn Glu Asp
025

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Lys Glu Phe Gly Lys Leu Glu Gly Gly Ala Ser Cys Ser Pro Ser Glu
1 5 10 15
Ser Asn Ala Ala Ser Ser Asn Ala Ile Cys Thr Ser Asn Gly Gly Glu
20 25 30
Thr Ile Gly Phe Val Asn Tyr Asn Lys Val Gly Thr Thr Thr Thr Leu
35 40 45
Glu Lys Arg Pro Glu Ile Leu Ile Phe Val Asn Gly Tyr Pro Ile Lys
50 55 60
Phe Leu Leu Asp Thr Gly Ala Asp Ile Thr Ile Leu Asn Arg Arg Asp
65 70 75 80
Phe Gln Val Lys Asn Ser Ile Glu Asn Gly Arg Gln Asn Met Ile Gly
85 90 95
Val Gly Gly Gly Lys Arg Gly Thr Asn Tyr Ile Asn Val His Leu Glu

```
            100                      105                      110
Ile Arg Asp Glu Asn Tyr Lys Thr Gln Cys Ile Phe Gly Asn Val Cys
        115                     120                     125
Val Leu Glu Asp Asn Ser Leu Ile Gln Pro Leu Leu Gly Arg Asp Asn
    130                     135                     140
Met Ile Lys Phe Asn Ile Arg Leu Val Met Ala Gln Ile Ser Asp Lys
145                     150                     155                 160
Ile Pro Val Val Lys Val Lys Met Lys Asp Pro Asn Lys Gly Pro Gln
                165                     170                     175
Ile Lys Gln Trp Pro Leu Thr Asn Glu Lys Ile Glu Ala Leu Thr Glu
            180                     185                     190
Ile Val Glu Arg Leu Glu Lys Glu Gly Lys Val Lys Arg Ala Asp Ser
        195                     200                     205
Asn Asn Pro Trp Asn Thr Pro Val Phe Ala Ile Lys Lys Lys Ser Gly
    210                     215                     220
Lys Trp Arg Met Leu Ile Asp Phe Arg Glu Leu Asn Lys Leu Thr Glu
225                     230                     235                 240
Lys Gly Ala Glu Val Gln Leu Gly Leu Pro His Pro Ala Gly Leu Gln
                245                     250                     255
Ile Lys Lys Gln Val Thr Val Leu Asp Ile Gly Asp Ala Tyr Phe Thr
            260                     265                     270
Ile Pro Leu Asp Pro Asp Tyr Ala Pro Tyr Thr Ala Phe Thr Leu Pro
        275                     280                     285
Arg Lys Asn Asn Ala Gly Pro Gly Arg Arg Phe Val Trp Cys Ser Leu
    290                     295                     300
Pro Gln Gly Trp Ile Leu Ser Pro Leu Ile Tyr Gln Ser Thr Leu Asp
305                     310                     315                 320
Asn Ile Ile Gln Pro Phe Ile Arg Gln Asn Pro Gln Leu Asp Ile Tyr
                325                     330                     335
Gln Tyr Met Asp Asp Ile Tyr Ile Gly Ser Asn Leu Ser Lys Lys Glu
            340                     345                     350
His Lys Glu Lys Val Glu Glu Leu Arg Lys Leu Leu Leu Trp Trp Gly
        355                     360                     365
Phe Glu Thr Pro Glu Asp Lys Leu Gln Glu Glu Pro Pro Tyr Thr Trp
    370                     375                     380
Met Gly Tyr Glu Leu His Pro Leu Thr Trp Thr Ile Gln Gln Lys Gln
385                     390                     395                 400
Leu Asp Ile Pro Glu Gln Pro Thr Leu Asn Glu Leu Gln Lys Leu Ala
                405                     410                     415
Gly Lys Ile Asn Trp Ala Ser Gln Ala Ile Pro Asp Leu Ser Ile Lys
            420                     425                     430
Ala Leu Thr Asn Met Met Arg Gly Asn Gln Asn Leu Asn Ser Thr Arg
        435                     440                     445
Gln Trp Thr Lys Glu Ala Arg Leu Glu Val Gln Lys Ala Lys Lys Ala
    450                     455                     460
Ile Glu Glu Gln Val Gln Leu Gly Tyr Tyr Asp Pro Ser Lys Glu Leu
465                     470                     475                 480
Tyr Ala Lys Leu Ser Leu Val Gly Pro His Gln Ile Ser Tyr Gln Val
                485                     490                     495
Tyr Gln Lys Asp Pro Glu Lys Ile Leu Trp Tyr Gly Lys Met Ser Arg
            500                     505                     510
Gln Lys Lys Lys Ala Glu Asn Thr Cys Asp Ile Ala Leu Arg Ala Cys
        515                     520                     525
Tyr Lys Ile Arg Glu Glu Ser Ile Ile Arg Ile Gly Lys Glu Pro Arg
    530                     535                     540
Tyr Glu Ile Pro Thr Ser Arg Glu Ala Trp Glu Ser Asn Leu Ile Asn
545                     550                     555                 560
Ser Pro Tyr Leu Lys Ala Pro Pro Pro Glu Val Glu Tyr Ile His Ala
                565                     570                     575
Ala Leu Asn Ile Lys Arg Ala Leu Ser Met Ile Lys Asp Ala Pro Ile
            580                     585                     590
Pro Gly Ala Glu Thr Trp Tyr Ile Asp Gly Gly Arg Lys Leu Gly Lys
        595                     600                     605
Ala Ala Lys Ala Ala Tyr Trp Thr Asp Thr Gly Lys Trp Arg Val Met
    610                     615                     620
Asp Leu Glu Gly Ser Asn Gln Lys Ala Glu Ile Gln Ala Leu Leu Leu
625                     630                     635                 640
Ala Leu Lys Ala Gly Ser Glu Glu Met Asn Ile Ile Thr Asp Ser Gln
                645                     650                     655
Tyr Val Ile Asn Ile Ile Leu Gln Gln Pro Asp Met Met Glu Gly Ile
```

```
                660                 665                 670
Trp Gln Glu Val Leu Glu Glu Leu Glu Lys Lys Thr Ala Ile Phe Ile
            675                 680                 685
Asp Trp Val Pro Gly His Lys Gly Ile Pro Gly Asn Glu Glu Val Asp
        690                 695                 700
Lys Leu Cys Gln Thr Met Met Ile Ile Glu Gly Asp Gly Ile Leu Asp
705                 710                 715                 720
Lys Arg Ser Glu Asp Ala Gly Tyr Asp Leu Leu Ala Ala Lys Glu Ile
                725                 730                 735
His Leu Leu Pro Gly Glu Val Lys Val Ile Pro Thr Gly Val Lys Leu
            740                 745                 750
Met Leu Pro Lys Gly Tyr Trp Gly Leu Ile Ile Gly Lys Ser Ser Ile
        755                 760                 765
Gly Ser Lys Gly Leu Asp Val Leu Gly Gly Val Ile Asp Glu Gly Tyr
    770                 775                 780
Arg Gly Glu Ile Gly Val Ile Met Ile Asn Val Ser Arg Lys Ser Ile
785                 790                 795                 800
Thr Leu Met Glu Arg Gln Lys Ile Ala Gln Leu Ile Ile Leu Pro Cys
                805                 810                 815
Lys His Glu Val Leu Glu Gln Gly Lys Val Val Met Asp Ser Glu Arg
            820                 825                 830
Gly Asp Asn Gly Tyr Gly Ser Thr Gly Val Phe Ser Ser Trp Val Asp
        835                 840                 845
Arg Ile Glu Glu Ala Glu Ile Asn His Glu Lys Phe His Ser Asp Pro
    850                 855                 860
Gln Tyr Leu Arg Thr Glu Phe Asn Leu Pro Lys Met Val Ala Glu Glu
865                 870                 875                 880
Ile Arg Arg Lys Cys Pro Val Cys Arg Ile Ile Gly Glu Gln Val Gly
                885                 890                 895
Gly Gln Leu Lys Ile Gly Pro Gly Ile Trp Gln Met Asp Cys Thr His
            900                 905                 910
Phe Asp Gly Lys Ile Ile Leu Val Gly Ile His Val Glu Ser Gly Tyr
        915                 920                 925
Ile Trp Ala Gln Ile Ile Ser Gln Glu Thr Ala Asp Cys Thr Val Lys
    930                 935                 940
Ala Val Leu Gln Leu Leu Ser Ala His Asn Val Thr Glu Leu Gln Thr
945                 950                 955                 960
Asp Asn Gly Pro Asn Phe Lys Asn Gln Lys Met Glu Gly Val Leu Asn
                965                 970                 975
Tyr Met Gly Val Lys His Lys Phe Gly Ile Pro Gly Asn Pro Gln Ser
            980                 985                 990
Gln Ala Leu Val Glu Asn Val Asn His Thr Leu Lys Val Trp Ile Gln
        995                1000                1005
Lys Phe Leu Pro Glu Thr Thr Ser Leu Asp Asn Ala Leu Ser Leu Ala
   1010                1015                1020
Val His Ser Leu Asn Phe Lys Arg Arg Gly Arg Ile Gly Gly Met Ala
025                1030                1035                1040
Pro Tyr Glu Leu Leu Ala Gln Gln Glu Ser Leu Arg Ile Gln Asp Tyr
               1045                1050                1055
Phe Ser Ala Ile Pro Gln Lys Leu Gln Ala Gln Trp Ile Tyr Tyr Lys
           1060                1065                1070
Asp Gln Lys Asp Lys Lys Trp Lys Gly Pro Met Arg Val Glu Tyr Trp
       1075                1080                1085
Gly Gln Gly Ser Val Leu Leu Lys Asp Glu Glu Lys Gly Tyr Phe Leu
   1090                1095                1100
Ile Pro Arg Arg His Ile Arg Arg Val Pro Glu Pro Cys Ala Leu Pro
105                1110                1115                1120
Glu Gly Asp Glu
               1
```

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 701 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

```
Met Glu Ala Val Ile Lys Val Ile Ser Ser Ala Cys Lys Thr Tyr Cys
 1               5                  10                  15
Gly Lys Thr Ser Pro Ser Lys Lys Glu Ile Gly Ala Met Leu Ser Leu
            20                  25                  30
Leu Gln Lys Glu Gly Leu Leu Met Ser Pro Ser Asp Leu Tyr Ser Pro
        35                  40                  45
Gly Ser Trp Asp Pro Ile Thr Ala Ala Leu Ser Gln Arg Ala Met Ile
    50                  55                  60
Leu Gly Lys Ser Gly Glu Leu Lys Thr Trp Gly Leu Val Leu Gly Ala
65                  70                  75                  80
Leu Lys Ala Ala Arg Glu Glu Gln Val Thr Ser Glu Gln Ala Lys Phe
                85                  90                  95
Trp Leu Gly Leu Gly Gly Gly Arg Val Ser Pro Pro Gly Pro Glu Cys
            100                 105                 110
Ile Glu Lys Pro Ala Thr Glu Arg Arg Ile Asp Lys Gly Glu Glu Val
        115                 120                 125
Gly Glu Thr Thr Val Gln Arg Asp Ala Lys Met Ala Pro Glu Glu Thr
    130                 135                 140
Ala Thr Pro Lys Thr Val Gly Thr Ser Cys Tyr His Cys Gly Thr Ala
145                 150                 155                 160
Ile Gly Cys Asn Cys Ala Thr Ala Ser Ala Pro Pro Pro Pro Tyr Val
                165                 170                 175
Gly Ser Gly Leu Tyr Pro Ser Leu Ala Gly Val Gly Glu Gln Gln Gly
            180                 185                 190
Gln Gly Gly Asp Thr Pro Pro Gly Ala Glu Gln Ser Arg Ala Glu Pro
        195                 200                 205
Gly His Ala Gly Gln Ala Pro Gly Pro Ala Leu Thr Asp Trp Ala Arg
    210                 215                 220
Val Arg Glu Glu Leu Ala Ser Thr Gly Pro Pro Val Val Ala Met Pro
225                 230                 235                 240
Val Val Ile Lys Thr Glu Gly Pro Ala Trp Thr Pro Leu Glu Pro Lys
                245                 250                 255
Leu Ile Thr Arg Leu Ala Asp Thr Val Arg Thr Lys Gly Leu Arg Ser
            260                 265                 270
Pro Ile Thr Met Ala Glu Val Glu Ala Leu Met Ser Ser Pro Leu Leu
        275                 280                 285
Pro His Asp Val Thr Asn Leu Met Arg Val Ile Leu Gly Pro Ala Pro
    290                 295                 300
Tyr Ala Leu Trp Met Asp Ala Trp Gly Val Gln Leu Gln Thr Val Ile
305                 310                 315                 320
Ala Ala Ala Thr Arg Asp Pro Arg His Pro Ala Asn Gly Gln Gly Arg
                325                 330                 335
Gly Glu Arg Thr Asn Leu Asn Arg Leu Lys Gly Leu Ala Asp Gly Met
            340                 345                 350
Val Gly Asn Pro Gln Gly Gln Ala Ala Leu Leu Arg Pro Gly Glu Leu
        355                 360                 365
Val Ala Ile Thr Ala Ser Ala Leu Gln Ala Phe Arg Glu Val Ala Arg
    370                 375                 380
Leu Ala Glu Pro Ala Gly Pro Trp Ala Asp Ile Met Gln Gly Pro Ser
385                 390                 395                 400
Glu Ser Phe Val Asp Phe Ala Asn Arg Leu Ile Lys Ala Val Glu Gly
                405                 410                 415
Ser Asp Leu Pro Pro Ser Ala Arg Ala Pro Val Ile Ile Asp Cys Phe
            420                 425                 430
Arg Gln Lys Ser Gln Pro Asp Ile Gln Gln Leu Ile Arg Thr Ala Pro
        435                 440                 445
Ser Thr Leu Thr Thr Pro Gly Glu Ile Ile Lys Tyr Val Leu Asp Arg
    450                 455                 460
Gln Lys Thr Ala Pro Leu Thr Asp Gln Gly Ile Ala Ala Ala Met Ser
465                 470                 475                 480
Ser Ala Ile Gln Pro Leu Ile Met Ala Val Val Asn Arg Glu Arg Asp
                485                 490                 495
Gly Gln Thr Gly Ser Gly Gly Arg Ala Arg Gly Leu Cys Tyr Thr Cys
            500                 505                 510
```

Gly Ser Pro Gly His Tyr Gln Ala Gln Cys Pro Lys Lys Arg Lys Ser
515 520 525
Gly Asn Ser Arg Glu Arg Cys Gln Leu Cys Asn Gly Met Gly His Asn
530 535 540
Ala Lys Gln Cys Arg Lys Arg Asp Gly Asn Gln Gly Gln Arg Pro Gly
545 550 555 560
Lys Gly Leu Ser Ser Gly Pro Trp Pro Gly Pro Glu Pro Pro Ala Val
565 570 575
Ser Leu Ala Met Thr Met Glu His Lys Asp Arg Pro Leu Val Arg Val
580 585 590
Ile Leu Thr Asn Thr Gly Ser His Pro Val Lys Gln Arg Ser Val Tyr
595 600 605
Ile Thr Ala Leu Leu Asp Ser Gly Ala Asp Ile Thr Ile Ile Ser Glu
610 615 620
Glu Asp Trp Pro Thr Asp Trp Pro Val Met Glu Ala Ala Asn Pro Gln
625 630 635 640
Ile His Gly Ile Gly Gly Gly Ile Pro Met Arg Lys Ser Arg Asp Met
645 650 655
Ile Glu Leu Gly Val Ile Asn Arg Asp Gly Ser Leu Glu Arg Pro Leu
660 665 670
Leu Leu Phe Pro Ala Val Ala Met Val Arg Gly Ser Ile Leu Gly Arg
675 680 685
Asp Cys Leu Gln Gly Leu Gly Leu Arg Leu Thr Asn Leu
690 695 700

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1199 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Gly Gly Gln Gly Gln Asp Pro Pro Pro Glu Pro Arg Ile Thr Leu Lys
1 5 10 15
Val Gly Gly Gln Pro Val Thr Phe Leu Val Asp Thr Gly Ala Gln His
20 25 30
Ser Val Leu Thr Gln Asn Pro Gly Pro Leu Ser Asp Lys Ser Ala Trp
35 40 45
Val Gln Gly Ala Thr Gly Gly Lys Arg Tyr Arg Trp Thr Thr Asp Arg
50 55 60
Lys Val His Leu Ala Thr Gly Lys Val Thr His Ser Phe Leu His Val
65 70 75 80
Pro Asp Cys Pro Tyr Pro Leu Leu Gly Arg Asp Leu Leu Thr Lys Leu
85 90 95
Lys Ala Gln Ile His Phe Glu Gly Ser Gly Ala Gln Val Met Gly Pro
100 105 110
Met Gly Gln Pro Leu Gln Val Leu Thr Leu Asn Ile Glu Asp Glu His
115 120 125
Arg Leu His Glu Thr Ser Lys Glu Pro Asp Val Ser Leu Gly Ser Thr
130 135 140
Trp Leu Ser Asp Phe Pro Gln Ala Trp Ala Glu Thr Gly Gly Met Gly
145 150 155 160
Leu Ala Val Arg Gln Ala Pro Leu Ile Ile Pro Leu Lys Ala Thr Ser
165 170 175
Thr Pro Val Ser Ile Lys Gln Tyr Pro Met Ser Gln Glu Ala Arg Leu
180 185 190
Gly Ile Lys Pro His Ile Gln Arg Leu Leu Asp Gln Gly Ile Leu Val
195 200 205
Pro Cys Gln Ser Pro Trp Asn Thr Pro Leu Leu Pro Val Lys Lys Pro
210 215 220
Gly Thr Asn Asp Tyr Arg Pro Val Gln Asp Leu Arg Glu Val Asn Lys

```
225                 230                 235                 240
Arg Val Glu Asp Ile His Pro Thr Val Pro Asn Pro Tyr Asn Leu Leu
                245                 250                 255
Ser Gly Leu Pro Pro Ser His Gln Trp Tyr Thr Val Leu Asp Leu Lys
            260                 265                 270
Asp Ala Phe Phe Cys Leu Arg Leu His Pro Thr Ser Gln Pro Leu Phe
        275                 280                 285
Ala Phe Glu Trp Arg Asp Pro Glu Met Gly Ile Ser Gly Gln Leu Thr
    290                 295                 300
Trp Thr Arg Leu Pro Gln Gly Phe Lys Asn Ser Pro Thr Leu Phe Asp
305                 310                 315                 320
Glu Ala Leu His Arg Asp Leu Ala Asp Phe Arg Ile Gln His Pro Asp
                325                 330                 335
Leu Ile Leu Leu Gln Tyr Val Asp Asp Leu Leu Leu Ala Ala Thr Ser
            340                 345                 350
Glu Leu Asp Cys Gln Gln Gly Thr Arg Ala Leu Leu Gln Thr Leu Gly
        355                 360                 365
Asn Leu Gly Tyr Arg Ala Ser Ala Lys Lys Ala Gln Ile Cys Gln Lys
    370                 375                 380
Gln Val Lys Tyr Leu Gly Tyr Leu Leu Lys Glu Gly Gln Arg Trp Leu
385                 390                 395                 400
Thr Glu Ala Arg Lys Glu Thr Val Met Gly Gln Pro Thr Pro Lys Thr
                405                 410                 415
Pro Arg Gln Leu Arg Glu Phe Leu Gly Thr Ala Gly Phe Cys Arg Leu
            420                 425                 430
Trp Ile Pro Gly Phe Ala Glu Met Ala Ala Pro Leu Tyr Pro Leu Thr
        435                 440                 445
Lys Thr Gly Thr Leu Phe Asn Trp Gly Pro Asp Gln Gln Lys Ala Tyr
    450                 455                 460
Gln Glu Ile Lys Gln Ala Leu Leu Thr Ala Pro Ala Leu Gly Leu Pro
465                 470                 475                 480
Asp Leu Thr Lys Pro Phe Glu Leu Phe Val Asp Glu Lys Gln Gly Tyr
                485                 490                 495
Ala Lys Gly Val Leu Thr Gln Lys Leu Gly Pro Trp Arg Arg Pro Val
            500                 505                 510
Ala Tyr Leu Ser Lys Lys Leu Asp Pro Val Ala Ala Gly Trp Pro Pro
        515                 520                 525
Cys Leu Arg Met Val Ala Ala Ile Ala Val Leu Thr Lys Asp Ala Gly
    530                 535                 540
Lys Leu Thr Met Gly Gln Pro Leu Val Ile Leu Ala Pro His Ala Val
545                 550                 555                 560
Glu Ala Leu Val Lys Gln Pro Pro Asp Arg Trp Leu Ser Asn Ala Arg
                565                 570                 575
Met Thr His Tyr Gln Ala Leu Leu Leu Asp Thr Asp Arg Val Gln Phe
            580                 585                 590
Gly Pro Val Val Ala Leu Asn Pro Ala Thr Leu Leu Pro Leu Pro Glu
        595                 600                 605
Glu Gly Leu Gln His Asn Cys Leu Asp Ile Leu Ala Glu Ala His Gly
    610                 615                 620
Thr Arg Pro Asp Leu Thr Asp Gln Pro Leu Pro Asp Ala Asp His Thr
625                 630                 635                 640
Trp Tyr Thr Asp Gly Ser Ser Leu Leu Gln Glu Gly Gln Arg Lys Ala
                645                 650                 655
Gly Ala Ala Val Thr Thr Glu Thr Glu Val Ile Trp Ala Lys Ala Leu
            660                 665                 670
Pro Ala Gly Thr Ser Ala Gln Arg Ala Glu Leu Ile Ala Leu Thr Gln
        675                 680                 685
Ala Leu Lys Met Ala Glu Gly Lys Lys Leu Asn Val Tyr Thr Asp Ser
    690                 695                 700
Arg Tyr Ala Phe Ala Thr Ala His Ile His Gly Glu Ile Tyr Arg Arg
705                 710                 715                 720
Arg Gly Leu Leu Thr Ser Glu Gly Lys Glu Ile Lys Asn Lys Asp Glu
                725                 730                 735
Ile Leu Ala Leu Leu Lys Ala Leu Phe Leu Pro Lys Arg Leu Ser Ile
            740                 745                 750
Ile His Cys Pro Gly His Gln Lys Gly His Ser Ala Glu Ala Arg Gly
        755                 760                 765
Asn Arg Met Ala Asp Gln Ala Ala Arg Lys Ala Ala Ile Thr Glu Thr
    770                 775                 780
Pro Asp Thr Ser Thr Leu Leu Ile Glu Asn Ser Ser Pro Tyr Thr Ser
```

785 790 795 800
Glu His Phe His Tyr Thr Val Thr Asp Ile Lys Asp Leu Thr Lys Leu
805 810 815
Gly Ala Ile Tyr Asp Lys Thr Lys Lys Tyr Trp Val Tyr Gln Gly Lys
820 825 830
Pro Val Met Pro Asp Gln Phe Thr Phe Glu Leu Leu Asp Phe Leu His
835 840 845
Gln Leu Thr His Leu Ser Phe Ser Lys Met Lys Ala Leu Leu Glu Arg
850 855 860
Ser His Ser Pro Tyr Tyr Met Leu Asn Arg Asp Arg Thr Leu Lys Asn
865 870 875 880
Ile Thr Glu Thr Cys Lys Ala Cys Ala Gln Val Asn Ala Ser Lys Ser
885 890 895
Ala Val Lys Gln Gly Thr Arg Val Arg Gly His Arg Pro Gly Thr His
900 905 910
Trp Glu Ile Asp Phe Thr Glu Ile Lys Pro Gly Leu Tyr Gly Tyr Lys
915 920 925
Tyr Leu Leu Val Phe Ile Asp Thr Phe Ser Gly Trp Ile Glu Ala Phe
930 935 940
Pro Thr Lys Lys Glu Thr Ala Lys Val Val Thr Lys Lys Leu Leu Glu
945 950 955 960
Glu Ile Phe Pro Arg Phe Gly Met Pro Gln Val Leu Gly Thr Asp Asn
965 970 975
Gly Pro Ala Phe Val Ser Lys Val Ser Gln Thr Val Ala Asp Leu Leu
980 985 990
Gly Ile Asp Trp Lys Leu His Cys Ala Tyr Arg Pro Gln Ser Ser Gly
995 1000 1005
Gln Val Glu Arg Met Asn Arg Thr Ile Lys Glu Thr Leu Thr Lys Leu
1010 1015 1020
Thr Leu Ala Thr Gly Ser Arg Asp Trp Val Leu Leu Leu Pro Leu Ala
025 1030 1035 1040
Leu Tyr Arg Ala Arg Asn Thr Pro Gly Pro His Gly Leu Thr Pro Tyr
1045 1050 1055
Glu Ile Leu Tyr Gly Ala Pro Pro Pro Leu Val Asn Phe Pro Asp Pro
1060 1065 1070
Asp Met Thr Arg Val Thr Asn Ser Pro Ser Leu Gln Ala His Leu Gln
1075 1080 1085
Ala Leu Tyr Leu Val Gln His Glu Val Trp Arg Pro Leu Ala Ala Ala
1090 1095 1100
Tyr Gln Glu Gln Leu Asp Arg Pro Val Val Pro His Pro Tyr Arg Val
105 1110 1115 1120
Gly Asp Thr Val Trp Val Arg Arg His Gln Thr Lys Asn Leu Glu Pro
1125 1130 1135
Arg Trp Lys Gly Pro Tyr Thr Val Leu Leu Thr Thr Pro Thr Ala Leu
1140 1145 1150
Lys Val Asp Gly Ile Ala Ala Trp Ile His Ala Ala His Val Lys Ala
1155 1160 1165
Ala Asp Pro Gly Gly Gly Pro Ser Ser Arg Leu Thr Trp Arg Val Gln
1170 1175 1180
Arg Ser Gln Asn Pro Leu Lys Ile Arg Leu Thr Arg Glu Ala Pro
185 1190 1195 1

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1204 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Thr Leu Asp Asp Gln Gly Gly Gln Gly Gln Glu Pro Pro Pro Glu Pro
1 5 10 15

Arg Ile Thr Leu Lys Val Gly Gly Gln Pro Val Thr Phe Leu Val Asp
20 25 30
Thr Gly Ala Gln His Ser Val Leu Thr Gln Asn Pro Gly Pro Leu Ser
35 40 45
Asp Lys Ser Ala Trp Val Gln Gly Ala Thr Gly Gly Lys Arg Tyr Arg
50 55 60
Trp Thr Thr Asp Arg Arg Val His Leu Ala Thr Gly Lys Val Thr His
65 70 75 80
Ser Phe Leu His Val Pro Asp Cys Pro Tyr Pro Leu Leu Gly Arg His
85 90 95
Leu Leu Thr Lys Leu Lys Ala Gln Ile His Phe Glu Gly Ser Gly Ala
100 105 110
Gln Val Val Gly Pro Met Gly Gln Pro Leu Gln Val Leu Thr Leu Asn
115 120 125
Ile Glu Asp Glu Tyr Arg Leu His Glu Thr Ser Lys Gly Pro Asp Val
130 135 140
Pro Leu Gly Ser Thr Trp Leu Ser Asp Phe Pro Gln Ala Trp Ala Glu
145 150 155 160
Thr Gly Gly Met Gly Leu Ala Phe Arg Gln Ala Pro Leu Ile Ile Ser
165 170 175
Leu Lys Ala Thr Ser Thr Pro Val Ser Ile Lys Gln Tyr Pro Met Ser
180 185 190
Gln Glu Ala Arg Leu Gly Ile Lys Pro His Ile Gln Arg Leu Leu Asp
195 200 205
Gln Gly Ile Leu Val Pro Cys Gln Ser Pro Trp Asn Thr Pro Leu Leu
210 215 220
Pro Val Lys Lys Pro Gly Thr Asn Asp Tyr Arg Pro Val Gln Asp Leu
225 230 235 240
Arg Glu Val Asn Lys Arg Val Glu Asp Ile His Pro Thr Val Pro Asn
245 250 255
Pro Tyr Asn Leu Leu Ser Gly Leu Pro Pro Ser His Gln Trp Tyr Thr
260 265 270
Val Leu Asp Leu Lys Asp Ala Phe Phe Cys Leu Arg Leu His Pro Thr
275 280 285
Ser Gln Ser Leu Phe Ala Phe Glu Trp Lys Asp Pro Glu Met Gly Ile
290 295 300
Ser Gly Gln Leu Thr Trp Thr Arg Leu Pro Gln Gly Phe Lys Asn Ser
305 310 315 320
Pro Thr Leu Phe Asp Glu Ala Leu His Arg Asp Leu Ala Asp Phe Arg
325 330 335
Ile Gln His Pro Asp Leu Ile Leu Leu Gln Tyr Val Asp Asp Leu Leu
340 345 350
Leu Ala Ala Thr Ser Glu Leu Asp Cys Gln Gln Gly Thr Arg Ala Leu
355 360 365
Leu Gln Thr Leu Gly Asp Leu Gly Tyr Arg Ala Ser Ala Lys Lys Ala
370 375 380
Gln Ile Cys Gln Lys Gln Val Lys Tyr Leu Gly Tyr Leu Leu Lys Glu
385 390 395 400
Gly Gln Arg Trp Leu Thr Glu Ala Arg Lys Glu Thr Val Met Gly Gln
405 410 415
Pro Thr Pro Lys Thr Pro Arg Gln Leu Arg Glu Phe Leu Gly Thr Ala
420 425 430
Gly Leu Cys Arg Leu Trp Ile Pro Gly Phe Ala Glu Met Ala Ala Pro
435 440 445
Leu Tyr Pro Leu Thr Lys Thr Gly Thr Leu Phe Lys Trp Gly Pro Asp
450 455 460
Gln Gln Lys Ala Tyr Gln Glu Ile Lys Gln Ala Leu Leu Thr Ala Pro
465 470 475 480
Ala Leu Gly Leu Pro Asp Leu Thr Lys Pro Phe Glu Leu Phe Val Asp
485 490 495
Glu Lys Gln Gly Tyr Ala Lys Gly Val Leu Thr Gln Lys Leu Gly Pro
500 505 510
Trp Arg Arg Pro Val Ala Tyr Leu Ser Lys Lys Leu Asp Pro Val Ala
515 520 525
Ala Gly Trp Pro Pro Cys Leu Arg Met Val Ala Ala Ile Ala Val Leu
530 535 540
Thr Lys Asp Val Gly Lys Leu Thr Met Gly Gln Pro Leu Val Ile Leu
545 550 555 560
Ala Pro His Ala Val Glu Ala Leu Val Lys Gln Pro Pro Asp Arg Trp
565 570 575

Leu Ser Asn Ala Arg Met Thr His Tyr Gln Ala Leu Leu Leu Asp Thr
580 585 590
Asp Arg Val Gln Phe Gly Pro Ile Val Ala Leu Asn Pro Ala Thr Leu
595 600 605
Leu Pro Leu Pro Glu Glu Gly Leu Gln His Asp Cys Leu Asp Ile Leu
610 615 620
Ala Glu Ala His Gly Thr Arg Pro Asp Leu Thr Asp Gln Pro Leu Pro
625 630 635 640
Asp Ala Asp His Thr Trp Tyr Thr Asp Gly Ser Ser Phe Leu Gln Glu
645 650 655
Gly Gln Arg Arg Ala Gly Ala Ala Val Thr Thr Glu Thr Glu Val Ile
660 665 670
Trp Ala Lys Ala Leu Pro Ala Gly Thr Ser Ala Gln Arg Ala Glu Leu
675 680 685
Ile Ala Leu Thr Gln Ala Leu Lys Met Ala Ala Gly Lys Lys Leu Asn
690 695 700
Val Tyr Thr Asp Ser Arg Tyr Ala Phe Ala Thr Ala His Ile His Gly
705 710 715 720
Glu Ile Tyr Arg Arg Arg Gly Leu Leu Thr Ser Glu Gly Lys Glu Ile
725 730 735
Lys Asn Lys Asp Glu Ile Leu Ala Leu Leu Lys Ala Leu Phe Leu Pro
740 745 750
Lys Arg Leu Ser Ile Ile His Cys Pro Gly His Gln Lys Gly Asn His
755 760 765
Ala Glu Ala Arg Gly Asn Arg Met Ala Asp Gln Ala Ala Arg Glu Val
770 775 780
Ala Thr Arg Glu Thr Pro Glu Thr Ser Thr Leu Leu Ile Glu Asn Ser
785 790 795 800
Ala Pro Tyr Thr Arg Glu His Phe His Tyr Thr Val Thr Asp Ile Lys
805 810 815
Asp Leu Thr Lys Leu Gly Ala Thr Tyr Asp Asp Ala Lys Lys Cys Trp
820 825 830
Val Tyr Gln Gly Lys Pro Val Met Pro Asp Gln Phe Thr Phe Glu Leu
835 840 845
Leu Asp Phe Leu His Gln Leu Thr His Leu Ser Phe Ser Lys Thr Lys
850 855 860
Ala Leu Leu Glu Arg Ser Tyr Ser Pro Ser Tyr Met Leu Asn Arg Asp
865 870 875 880
Arg Thr Leu Lys Asp Ile Thr Glu Thr Cys Lys Ala Cys Ala Gln Val
885 890 895
Asn Ala Ser Lys Ser Ala Val Lys Gln Gly Thr Arg Val Arg Gly His
900 905 910
Arg Pro Gly Thr His Trp Glu Ile Asp Phe Thr Glu Val Lys Pro Gly
915 920 925
Leu Tyr Gly Tyr Lys Tyr Leu Leu Val Phe Val Asp Thr Phe Ser Gly
930 935 940
Trp Val Glu Ala Phe Pro Thr Lys Lys Glu Thr Ala Lys Val Val Thr
945 950 955 960
Lys Lys Leu Leu Glu Glu Ile Phe Pro Arg Phe Gly Met Pro Gln Val
965 970 975
Leu Gly Thr Asp Asn Gly Pro Ala Phe Val Ser Lys Val Ser Gln Thr
980 985 990
Val Ala Asp Leu Leu Gly Val Asp Trp Lys Leu His Cys Ala Tyr Arg
995 1000 1005
Pro Gln Ser Ser Gly Gln Val Glu Arg Met Asn Arg Thr Ile Lys Glu
1010 1015 1020
Thr Leu Thr Lys Leu Thr Leu Ala Thr Gly Ser Arg Asp Trp Val Leu
025 1030 1035 1040
Leu Leu Pro Leu Ala Leu Tyr Arg Ala Arg Asn Thr Pro Gly Pro His
1045 1050 1055
Gly Leu Thr Pro Tyr Glu Ile Leu Tyr Gly Ala Pro Pro Pro Leu Val
1060 1065 1070
Asn Phe Pro Asp Pro Asp Met Ala Lys Val Thr His Asn Pro Ser Leu
1075 1080 1085
Gln Ala His Leu Gln Ala Leu Tyr Leu Val Gln His Glu Val Trp Arg
1090 1095 1100
Pro Leu Ala Ala Ala Tyr Gln Glu Gln Leu Asp Arg Pro Val Val Pro
105 1110 1115 1120
His Pro Phe Arg Val Gly Asp Thr Val Trp Val Arg Arg His Gln Thr
1125 1130 1135

Lys Asn Leu Glu Pro Arg Trp Lys Gly Pro Tyr Thr Val Leu Leu Thr
1140 1145 1150
Thr Pro Thr Ala Leu Lys Val Asp Gly Ile Ala Ala Trp Ile His Ala
1155 1160 1165
Ala His Val Lys Ala Ala Asp Thr Arg Ile Glu Pro Pro Ala Glu Ser
1170 1175 1180
Thr Trp Arg Val Gln Arg Ser Gln Asn Pro Leu Lys Ile Arg Leu Thr
185 1190 1195 1200
Arg Gly Thr Ser
1

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 340 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Met Asp Ser Ala Ala Pro Ala Leu Ser Pro Ala Leu Thr Ala Leu Thr
1 5 10 15
Asp Gln Ser Ala Thr Ala Asp Leu Ala Ile Gln Ile Pro Lys Cys Pro
20 25 30
Asp Pro Glu Arg Tyr Phe Tyr Thr Ser Gln Cys Pro Asp Ile Asn His
35 40 45
Leu Arg Ser Leu Ser Ile Leu Asn Arg Trp Leu Glu Thr Glu Leu Val
50 55 60
Phe Val Gly Asp Glu Glu Asp Val Ser Lys Leu Ser Glu Gly Glu Leu
65 70 75 80
Ser Phe Tyr Arg Phe Leu Phe Ala Phe Leu Ser Ala Ala Asp Asp Leu
85 90 95
Val Thr Glu Asn Leu Gly Gly Leu Ser Gly Leu Phe Glu Gln Lys Asp
100 105 110
Ile Leu His Tyr Tyr Val Glu Gln Glu Cys Ile Glu Val Val His Ser
115 120 125
Arg Val Tyr Asn Ile Ile Gln Leu Val Leu Phe His Asn Asn Asp Gln
130 135 140
Ala Arg Arg Glu Tyr Val Ala Gly Thr Ile Asn His Pro Ala Ile Arg
145 150 155 160
Ala Lys Val Asp Trp Leu Glu Ala Arg Val Arg Glu Cys Ala Ser Val
165 170 175
Pro Glu Lys Phe Ile Leu Met Ile Leu Ile Glu Gly Ile Phe Phe Ala
180 185 190
Ala Ser Phe Ala Ala Ile Ala Tyr Leu Arg Thr Asn Asn Leu Leu Arg
195 200 205
Val Thr Cys Gln Ser Asn Asp Leu Ile Ser Arg Asp Glu Ala Val His
210 215 220
Thr Thr Ala Ser Cys Tyr Ile Tyr Asn Asn Tyr Leu Gly Gly His Ala
225 230 235 240
Lys Pro Pro Pro Asp Arg Val Tyr Gly Leu Phe Arg Gln Ala Val Glu
245 250 255
Ile Glu Ile Gly Phe Ile Arg Ser Gln Ala Pro Thr Asp Ser His Ile
260 265 270
Leu Ser Pro Ala Ala Leu Ala Ala Ile Glu Asn Tyr Val Arg Ph Ser
275 280 285
Ala Asp Arg Leu Leu Gly Leu Ile His Met Lys Pro Leu Phe Ser Ala
290 295 300
Pro Pro Pro Asp Ala Ser Phe Pro Leu Ser Leu Met Ser Thr Asp Lys
305 310 315 320
His Thr Asn Phe Phe Glu Cys Arg Ser Thr Ser Tyr Ala Gly Ala Val
325 330 335
Val Asn Asp Leu

340

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

```
Met Asp Pro Ala Val Ser Pro Ala Ser Thr Asp Pro Leu Asp Thr His
 1               5                  10                  15
Ala Ser Gly Ala Gly Ala Ala Pro Ile Pro Val Cys Pro Thr Pro Glu
            20                  25                  30
Arg Tyr Phe Tyr Thr Ser Gln Cys Pro Asp Ile Asn His Leu Arg Ser
        35                  40                  45
Leu Ser Ile Leu Asn Arg Trp Leu Glu Thr Glu Leu Val Phe Val Gly
    50                  55                  60
Asp Glu Glu Asp Val Ser Lys Leu Ser Glu Gly Glu Leu Gly Phe Tyr
65                  70                  75                  80
Arg Phe Leu Phe Ala Phe Leu Ser Ala Ala Asp Asp Leu Val Thr Glu
                85                  90                  95
Asn Leu Gly Gly Leu Ser Gly Leu Phe Glu Gln Lys Asp Ile Leu His
            100                 105                 110
Tyr Tyr Val Glu Gln Glu Cys Ile Glu Val Val His Ser Arg Val Tyr
        115                 120                 125
Asn Ile Ile Gln Leu Val Leu Phe His Asn Asn Asp Gln Ala Arg Arg
    130                 135                 140
Ala Tyr Val Ala Arg Thr Ile Asn His Pro Ala Ile Arg Val Lys Val
145                 150                 155                 160
Asp Trp Leu Glu Ala Arg Val Arg Glu Cys Asp Ser Ile Pro Glu Lys
                165                 170                 175
Phe Ile Leu Met Ile Leu Ile Glu Gly Val Phe Phe Ala Ala Ser Phe
            180                 185                 190
Ala Ala Ile Ala Tyr Leu Arg Thr Asn Asn Leu Leu Arg Val Thr Cys
        195                 200                 205
Gln Ser Asn Asp Leu Ile Ser Arg Asp Glu Ala Val His Thr Thr Ala
    210                 215                 220
Ser Cys Tyr Ile Tyr Asn Asn Tyr Leu Gly Gly His Ala Lys Pro Glu
225                 230                 235                 240
Ala Ala Arg Val Tyr Arg Leu Phe Arg Glu Ala Val Asp Ile Glu Ile
                245                 250                 255
Gly Phe Ile Arg Ser Gln Ala Pro Thr Asp Ser Ser Ile Leu Ser Pro
            260                 265                 270
Gly Ala Leu Ala Ala Ile Glu Asn Tyr Val Arg Phe Ser Ala Asp Arg
        275                 280                 285
Leu Leu Gly Leu Ile His Met Gln Pro Leu Tyr Ser Ala Pro Ala Pro
    290                 295                 300
Asp Ala Ser Phe Pro Leu Ser Leu Met Ser Thr Asp Lys His Thr Asn
305                 310                 315                 320
Phe Phe Glu Cys Arg Ser Thr Ser Tyr Ala Gly Ala Val Val Asn Asp
                325                 330                 335
Leu
```

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 302 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```
Met Ser Lys Leu Leu Tyr Val Arg Asp His Glu Gly Phe Ala Cys Leu
 1               5                  10                  15
Thr Val Glu Thr His Arg Asn Arg Trp Phe Ala Ala His Ile Val Leu
            20                  25                  30
Thr Lys Asp Cys Gly Cys Leu Lys Leu Leu Asn Glu Arg Asp Leu Glu
        35                  40                  45
Phe Tyr Lys Phe Leu Phe Thr Phe Leu Ala Met Ala Glu Lys Leu Val
    50                  55                  60
Asn Phe Asn Ile Asp Glu Leu Val Thr Ser Phe Glu Ser His Asp Ile
65                  70                  75                  80
Asp His Tyr Tyr Thr Glu Gln Lys Ala Met Glu Asn Val His Gly Glu
                85                  90                  95
Thr Tyr Ala Asn Ile Leu Asn Met Leu Phe Asp Gly Asp Arg Ala Ala
            100                 105                 110
Met Asn Ala Tyr Ala Glu Ala Ile Met Ala Asp Glu Ala Leu Gln Ala
        115                 120                 125
Lys Ile Ser Trp Leu Arg Asp Lys Val Ala Ala Ala Val Thr Leu Pro
    130                 135                 140
Glu Lys Ile Leu Val Phe Leu Leu Ile Glu Gly Ile Phe Phe Ile Ser
145                 150                 155                 160
Ser Phe Tyr Ser Ile Ala Leu Leu Arg Val Arg Gly Leu Met Pro Gly
                165                 170                 175
Ile Cys Leu Ala Asn Asn Tyr Ile Ser Arg Asp Glu Leu Leu His Thr
            180                 185                 190
Arg Ala Ala Ser Leu Leu Tyr Asn Ser Met Thr Ala Lys Ala Asp Arg
        195                 200                 205
Pro Arg Ala Thr Trp Ile Gln Glu Leu Phe Arg Thr Ala Val Glu Val
    210                 215                 220
Glu Thr Ala Phe Ile Glu Ala Arg Gly Glu Gly Val Thr Leu Val Asp
225                 230                 235                 240
Val Arg Ala Ile Lys Gln Phe Leu Glu Ala Thr Ala Asp Arg Ile Leu
                245                 250                 255
Gly Asp Ile Gly Gln Ala Pro Leu Tyr Gly Thr Pro Pro Pro Lys Asp
            260                 265                 270
Cys Pro Leu Thr Tyr Met Thr Ser Ile Lys Gln Thr Asn Phe Phe Glu
        275                 280                 285
Gln Glu Ser Ser Asp Tyr Thr Met Leu Val Val Asp Asp Leu
    290                 295                 300
```

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 389 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

```
Met Leu Ser Leu Arg Val Pro Leu Ala Pro Ile Thr Asp Pro Gln Gln
 1               5                  10                  15
Leu Gln Leu Ser Pro Leu Lys Gly Leu Ser Leu Val Asp Lys Glu Asn
            20                  25                  30
Thr Pro Pro Ala Leu Ser Gly Thr Arg Val Leu Ala Ser Lys Thr Ala
        35                  40                  45
Arg Arg Ile Phe Gln Glu Pro Thr Glu Pro Lys Thr Lys Ala Ala Ala
```

50 55 60
Pro Gly Val Glu Asp Glu Pro Leu Leu Arg Glu Asn Pro Arg Arg Phe
65 70 75 80
Val Ile Phe Pro Ile Glu Tyr His Asp Ile Trp Gln Met Tyr Lys Lys
85 90 95
Ala Glu Ala Ser Phe Trp Thr Ala Glu Glu Val Asp Leu Ser Lys Asp
100 105 110
Ile Gln His Trp Glu Ser Leu Lys Pro Glu Glu Arg Tyr Phe Ile Ser
115 120 125
His Val Leu Ala Phe Phe Ala Ala Ser Asp Gly Ile Val Asn Glu Asn
130 135 140
Leu Val Glu Arg Phe Ser Gln Glu Val Gln Ile Thr Glu Ala Arg Cys
145 150 155 160
Phe Tyr Gly Phe Gln Ile Ala Met Glu Asn Ile His Ser Glu Met Tyr
165 170 175
Ser Leu Leu Ile Asp Thr Tyr Ile Lys Asp Pro Lys Glu Arg Glu Phe
180 185 190
Leu Phe Asn Ala Ile Glu Thr Met Pro Cys Val Lys Lys Lys Ala Asp
195 200 205
Trp Ala Leu Arg Trp Ile Gly Asp Lys Glu Ala Thr Tyr Gly Glu Arg
210 215 220
Val Val Ala Phe Ala Ala Val Glu Gly Ile Phe Phe Ser Gly Ser Phe
225 230 235 240
Ala Ser Ile Phe Trp Leu Lys Lys Arg Gly Leu Met Pro Gly Leu Thr
245 250 255
Phe Ser Asn Glu Leu Ile Ser Arg Asp Glu Gly Leu His Cys Asp Phe
260 265 270
Ala Cys Leu Met Phe Lys His Leu Val His Lys Pro Ser Glu Glu Arg
275 280 285
Val Arg Glu Ile Ile Ile Asn Ala Val Arg Ile Glu Gln Glu Phe Leu
290 295 300
Thr Glu Ala Leu Pro Val Lys Leu Ile Gly Met Asn Cys Thr Leu Met
305 310 315 320
Lys Gln Tyr Ile Glu Phe Val Ala Asp Arg Leu Met Leu Glu Leu Gly
325 330 335
Phe Ser Lys Val Phe Arg Val Glu Asn Pro Phe Asp Phe Met Glu Asn
340 345 350
Ile Ser Leu Glu Gly Lys Thr Asn Phe Phe Glu Lys Arg Val Gly Glu
355 360 365
Tyr Gln Arg Met Gly Val Met Ser Ser Pro Thr Glu Asn Ser Phe Thr
370 375 380
Leu Asp Ala Asp Phe
385

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 319 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Glu Pro Ile Leu Ala Pro Asn Pro Asn Arg Phe Val Ile Phe Pro
1 5 10 15
Ile Gln Tyr Tyr Asp Ile Trp Asn Met Tyr Lys Lys Ala Glu Ala Ser
20 25 30
Phe Trp Thr Val Glu Glu Val Asp Ile Ser Lys Asp Ile Asn Asp Trp
35 40 45
Asn Lys Leu Thr Pro Asp Glu Lys Tyr Phe Ile Lys His Val Leu Ala
50 55 60
Phe Phe Ala Ala Ser Asp Gly Ile Val Asn Glu Asn Leu Ala Glu Arg
65 70 75 80

Phe Cys Thr Glu Val Gln Ile Thr Glu Ala Arg Cys Phe Tyr Gly Phe
85 90 95
Gln Met Ala Ile Glu Asn Ile His Ser Glu Met Tyr Ser Leu Leu Ile
100 105 110
Asp Thr Tyr Val Lys Asp Ser Asn Glu Lys Asn Tyr Leu Phe Asn Ala
115 120 125
Ile Glu Thr Met Pro Cys Val Lys Lys Lys Ala Asp Trp Ala Gln Lys
130 135 140
Trp Ile His Asp Ser Ala Gly Tyr Gly Glu Arg Leu Ile Ala Phe Ala
145 150 155 160
Ala Val Glu Gly Ile Phe Phe Ser Gly Ser Phe Ala Ser Ile Phe Trp
165 170 175
Leu Lys Lys Arg Gly Leu Met Pro Gly Leu Thr Phe Ser Asn Glu Leu
180 185 190
Ile Ser Arg Asp Glu Gly Leu His Cys Asp Phe Ala Cys Leu Met Phe
195 200 205
Lys His Leu Leu His Pro Pro Ser Glu Glu Thr Val Arg Ser Ile Ile
210 215 220
Thr Asp Ala Val Ser Ile Glu Gln Glu Phe Leu Thr Ala Ala Leu Pro
225 230 235 240
Val Lys Leu Ile Gly Met Asn Cys Glu Met Met Lys Thr Tyr Ile Glu
245 250 255
Phe Val Ala Asp Arg Leu Ile Ser Glu Leu Gly Phe Lys Lys Ile Tyr
260 265 270
Asn Val Thr Asn Pro Phe Asp Phe Met Glu Asn Ile Ser Leu Glu Gly
275 280 285
Lys Thr Asn Phe Phe Glu Lys Arg Val Gly Glu Tyr Gln Lys Met Gly
290 295 300
Val Met Ser Gln Glu Asp Asn His Phe Ser Leu Asp Val Asp Phe
305 310 315

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 390 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Met Leu Ser Val Arg Thr Pro Leu Ala Thr Ile Ala Asp Gln Gln Gln
1 5 10 15
Leu Gln Leu Ser Pro Leu Lys Arg Leu Thr Leu Ala Asp Lys Glu Asn
20 25 30
Thr Pro Pro Thr Leu Ser Ser Thr Arg Val Leu Ala Ser Lys Ala Ala
35 40 45
Arg Arg Ile Phe Gln Asp Ser Ala Glu Leu Glu Ser Lys Ala Pro Thr
50 55 60
Asn Pro Ser Val Glu Asp Glu Pro Leu Leu Arg Glu Asn Pro Arg Arg
65 70 75 80
Phe Val Val Phe Pro Ile Glu Tyr His Asp Ile Trp Gln Met Tyr Lys
85 90 95
Lys Ala Glu Ala Ser Phe Trp Thr Ala Glu Glu Val Asp Leu Ser Lys
100 105 110
Asp Ile Gln His Trp Glu Ala Leu Lys Pro Asp Glu Arg His Phe Ile
115 120 125
Ser His Val Leu Ala Phe Phe Ala Ala Ser Asp Gly Ile Val Asn Glu
130 135 140
Asn Leu Val Glu Arg Phe Ser Gln Glu Val Gln Val Thr Glu Ala Arg
145 150 155 160
Cys Phe Tyr Gly Phe Gln Ile Ala Met Glu Asn Ile His Ser Glu Met
165 170 175
Tyr Ser Leu Leu Ile Asp Thr Tyr Ile Lys Asp Pro Lys Glu Arg Glu

```
                180                     185                     190
Tyr Leu Phe Asn Ala Ile Glu Thr Met Pro Cys Val Lys Lys Lys Ala
            195                     200                     205
Asp Trp Ala Leu Arg Trp Ile Gly Asp Lys Glu Ala Thr Tyr Gly Glu
        210                     215                     220
Arg Val Val Ala Phe Ala Ala Val Glu Gly Ile Phe Phe Ser Gly Ser
225                     230                     235                     240
Phe Ala Ser Ile Phe Trp Leu Lys Lys Arg Gly Leu Met Pro Gly Leu
                245                     250                     255
Thr Phe Ser Asn Glu Leu Ile Ser Arg Asp Glu Gly Leu His Cys Asp
            260                     265                     270
Phe Ala Cys Leu Met Phe Lys His Leu Val His Lys Pro Ala Glu Gln
        275                     280                     285
Arg Val Arg Glu Ile Ile Thr Asn Ala Val Arg Ile Glu Gln Glu Phe
    290                     295                     300
Leu Thr Glu Ala Leu Pro Val Lys Leu Ile Gly Met Asn Cys Thr Leu
305                     310                     315                     320
Met Lys Gln Tyr Ile Glu Phe Val Ala Asp Arg Leu Met Leu Glu Leu
                325                     330                     335
Gly Phe Asn Lys Ile Phe Arg Val Glu Asn Pro Phe Asp Phe Met Glu
            340                     345                     350
Asn Ile Ser Leu Glu Gly Lys Thr Asn Phe Phe Glu Lys Arg Val Gly
        355                     360                     365
Glu Tyr Gln Arg Met Gly Val Met Ser Asn Ser Thr Glu Asn Ser Phe
    370                     375                     380
Thr Leu Asp Ala Asp Phe
385                 390
```

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 399 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
Met Pro Lys Glu Thr Pro Ser Lys Ala Ala Ala Asp Ala Leu Ser Asp
1               5                   10                  15
Leu Glu Ile Lys Asp Ser Lys Ser Asn Leu Asn Lys Glu Leu Glu Thr
            20                  25                  30
Leu Arg Glu Glu Asn Arg Val Lys Ser Asp Met Leu Lys Glu Lys Leu
        35                  40                  45
Ser Lys Asp Ala Glu Asn His Lys Ala Tyr Leu Lys Ser His Gln Val
    50                  55                  60
His Arg His Lys Leu Lys Glu Met Glu Lys Glu Glu Pro Leu Leu Asn
65                  70                  75                  80
Glu Asp Lys Glu Arg Thr Val Leu Phe Pro Ile Lys Tyr His Glu Ile
                85                  90                  95
Trp Gln Ala Tyr Lys Arg Ala Glu Ala Ser Phe Trp Thr Ala Glu Glu
            100                 105                 110
Ile Asp Leu Ser Lys Asp Ile His Asp Trp Asn Asn Arg Met Asn Glu
        115                 120                 125
Asn Glu Arg Phe Phe Ile Ser Arg Val Leu Ala Phe Phe Ala Ala Ser
    130                 135                 140
Asp Gly Ile Val Asn Glu Asn Leu Val Glu Asn Phe Ser Thr Glu Val
145                 150                 155                 160
Gln Ile Pro Glu Ala Lys Ser Phe Tyr Gly Phe Gln Ile Met Ile Glu
                165                 170                 175
Asn Ile His Ser Glu Thr Tyr Ser Leu Leu Ile Asp Thr Tyr Ile Lys
            180                 185                 190
Asp Pro Lys Glu Ser Glu Phe Leu Phe Asn Ala Ile His Thr Ile Pro
        195                 200                 205
```

Glu Ile Gly Glu Lys Ala Glu Trp Ala Leu Arg Trp Ile Gln Asp Ala
210 215 220
Asp Ala Leu Phe Gly Glu Arg Leu Val Ala Phe Ala S r Ile Glu Gly
225 230 235 240
Val Phe Phe Ser Gly Ser Phe Ala Ser Ile Phe Trp Leu Lys Lys Arg
245 250 255
Gly Met Met Pro Gly Leu Thr Phe Ser Asn Glu Leu Ile Cys Arg Asp
260 265 270
Glu Gly Leu His Thr Asp Phe Ala Cys Leu Leu Phe Ala His Leu Lys
275 280 285
Asn Lys Pro Asp Pro Ala Ile Val Glu Lys Ile Val Thr Glu Ala Val
290 295 300
Glu Ile Glu Gln Arg Tyr Phe Leu Asp Ala Leu Pro Val Ala Leu Leu
305 310 315 320
Gly Met Asn Ala Asp Leu Met Asn Gln Tyr Val Glu Phe Val Ala Asp
325 330 335
Arg Leu Leu Val Ala Phe Gly Asn Lys Lys Tyr Tyr Lys Val Glu Asn
340 345 350
Pro Phe Asp Phe Met Glu Asn Ile Ser Leu Ala Gly Lys Thr Asn Phe
355 360 365
Phe Glu Lys Arg Val Ser Asp Tyr Gln Lys Ala Gly Val Met Ser Lys
370 375 380
Ser Thr Lys Gln Glu Ala Gly Ala Phe Thr Phe Asn Glu Asp Phe
385 390 395

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 375 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Ala Tyr Thr Thr Phe Ser Gln Thr Lys Asn Asp Gln Leu Lys Glu Pro
1 5 10 15
Met Phe Phe Gly Gln Pro Val Asn Val Ala Arg Tyr Asp Gln Gln Lys
20 25 30
Tyr Asp Ile Phe Glu Lys Leu Ile Glu Lys Gln Leu Ser Phe Phe Trp
35 40 45
Arg Pro Glu Glu Val Asp Val Ser Arg Asp Arg Ile Asp Tyr Gln Ala
50 55 60
Leu Pro Glu His Glu Lys His Ile Phe Ile Ser Asn Leu Lys Tyr Gln
65 70 75 80
Thr Leu Leu Asp Ser Ile Gln Gly Arg Ser Pro Asn Val Ala Leu Leu
85 90 95
Pro Leu Ile Ser Ile Pro Glu Leu Glu Thr Trp Val Glu Thr Trp Ala
100 105 110
Phe Ser Glu Thr Ile His Ser Arg Ser Tyr Thr His Ile Ile Arg Asn
115 120 125
Ile Val Asn Asp Pro Ser Val Val Phe Asp Asp Ile Val Thr Asn Glu
130 135 140
Gln Ile Gln Lys Arg Ala Glu Gly Ile Ser Ser Tyr Tyr Asp Glu Leu
145 150 155 160
Ile Glu Met Thr Ser Tyr Trp His Leu Leu Gly Glu Gly Thr His Thr
165 170 175
Val Asn Gly Lys Thr Val Thr Val Ser Leu Arg Glu Leu Lys Lys Lys
180 185 190
Leu Tyr Leu Cys Leu Met Ser Val Asn Ala Leu Glu Ala Ile Arg Phe
195 200 205
Tyr Val Ser Phe Ala Cys Ser Phe Ala Phe Ala Glu Arg Glu Leu Met
210 215 220
Glu Gly Asn Ala Lys Ile Ile Arg Leu Ile Ala Arg Asp Glu Ala Leu

225 230 235 240
His Leu Thr Gly Thr Gln His Met Leu Asn Leu Leu Arg Ser Gly Ala
245 250 255
Asp Asp Pro Glu Met Ala Glu Ile Ala Glu Glu Cys Lys Gln Glu Cys
260 265 270
Tyr Asp Leu Phe Val Gln Ala Ala Gln Gln Glu Lys Asp Trp Ala Asp
275 280 285
Tyr Leu Phe Arg Asp Gly Ser Met Ile Gly Leu Asn Lys Asp Ile Leu
290 295 300
Cys Gln Tyr Val Glu Tyr Ile Thr Asn Ile Arg Met Gln Ala Val Gly
305 310 315 320
Leu Asp Leu Pro Phe Gln Thr Arg Ser Asn Pro Ile Pro Trp Ile Asn
325 330 335
Thr Trp Leu Val Ser Asp Asn Val Gln Val Ala Pro Gln Glu Val Glu
340 345 350
Val Ser Ser Tyr Leu Val Gly Gln Ile Asp Ser Glu Val Asp Thr Asp
355 360 365
Asp Leu Ser Asn Phe Gln Leu
370 375

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 375 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

Ala Tyr Thr Thr Phe Ser Gln Asn Lys Asn Asp Gln Leu Lys Glu Pro
1 5 10 15
Met Phe Phe Gly Gln Asn Val Asn Val Ala Arg Tyr Asp Gln Gln Lys
20 25 30
Tyr Glu Thr Phe Glu Lys Leu Ile Glu Lys Gln Leu Ser Phe Phe Trp
35 40 45
Arg Pro Glu Glu Val Asp Val Ser Gln Asp Arg Ile Asp Tyr Ala Ala
50 55 60
Leu Pro Glu His Glu Lys His Ile Phe Ile Ser Asn Leu Lys Tyr Gln
65 70 75 80
Thr Leu Leu Asp Ser Ile Gln Gly Arg Ser Pro Asn Val Ala Leu Leu
85 90 95
Pro Leu Val Ser Ile Pro Glu Leu Glu Thr Trp Ile Glu Thr Trp Thr
100 105 110
Phe Ser Glu Thr Ile His Ser Arg Ser Tyr Thr His Ile Ile Arg Asn
115 120 125
Ile Val Asn Asp Pro Ser Ile Val Phe Asp Asp Ile Val Thr Asn Glu
130 135 140
Glu Ile Ile Lys Arg Ala Gln Asp Ile Ser Ser Tyr Tyr Asp Asp Leu
145 150 155 160
Ile Arg Asp Ser Gln Leu Tyr Gly Leu Tyr Gly Glu Gly Thr Tyr Thr
165 170 175
Val Asp Gly Lys Glu Cys Val Val Thr Leu Arg Ser Leu Lys Lys Gln
180 185 190
Leu Tyr Leu Cys Leu Met Ser Val Asn Ala Leu Glu Ala Ile Arg Phe
195 200 205
Tyr Val Ser Phe Ala Cys Ser Phe Ala Phe Ala Glu Arg Arg Leu Met
210 215 220
Glu Gly Asn Ala Lys Ile Il Lys Ph Ile Ala Arg Asp Glu Ala Leu
225 230 235 240
His Leu Thr Gly Thr Gln His Ile Leu Asn Ile Met Ala Ala Gly Gln
245 250 255
Asp Asp Pro Glu Met Ala Glu Ile Ala Glu Glu Cys Lys Gln Glu Ala
260 265 270

```
Tyr Asp Leu Phe Val Ala Ala Ala Glu Gln Glu Lys Ala Trp Ala Asp
        275                 280                 285
Tyr Leu Phe Lys Asp Gly Ser Met Ile Gly Leu Asn Arg Asp Ile Leu
    290                 295                 300
Val Gln Tyr Val Glu Tyr Ile Thr Asn Ile Arg Met Gln Ala Val Gly
305                 310                 315                 320
Leu Pro Leu Pro Phe Gln Thr Arg Ser Asn Pro Ile Pro Trp Ile Asn
                325                 330                 335
Ala Trp Leu Val Ser Asp Asn Val Gln Val Ala Pro Gln Glu Val Glu
            340                 345                 350
Val Ser Ser Tyr Leu Val Gly Gln Ile Asp Ser Lys Val Asp Thr Asn
        355                 360                 365
Asp Phe Asp Asp Phe Ser Leu
    370                 375
```

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 479 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Met Asp Ala Asp Gly Ala Ser Pro Pro Pro Pro Arg Pro Ala Gly Gly
1               5                   10                  15
Pro Lys Asn Thr Pro Ala Ala Pro Pro Leu Tyr Ala Thr Gly Arg Leu
            20                  25                  30
Ser Gln Ala Gln Leu Met Pro Ser Pro Pro Met Pro Val Pro Pro Ala
        35                  40                  45
Ala Leu Phe Asn Arg Leu Leu Asp Asp Leu Gly Phe Ser Ala Gly Pro
    50                  55                  60
Ala Leu Cys Thr Met Leu Asp Thr Trp Asn Glu Asp Leu Phe Ser Ala
65                  70                  75                  80
Leu Pro Thr Asn Ala Asp Leu Tyr Arg Glu Cys Lys Phe Leu Ser Thr
                85                  90                  95
Leu Pro Ser Asp Val Val Glu Trp Gly Asp Ala Tyr Val Pro Glu Arg
            100                 105                 110
Ala Gln Ile Asp Ile Arg Ala His Gly Asp Val Ala Phe Pro Thr Leu
        115                 120                 125
Pro Ala Thr Arg Asp Gly Leu Gly Leu Tyr Tyr Glu Ala Leu Ser Arg
    130                 135                 140
Phe Phe His Ala Glu Leu Arg Ala Arg Glu Glu Ser Tyr Arg Thr Val
145                 150                 155                 160
Leu Ala Asn Phe Cys Ser Ala Leu Tyr Arg Tyr Leu Arg Ala Ser Val
                165                 170                 175
Arg Gln Leu His Arg Gln Ala His Met Arg Gly Arg Asp Arg Asp Leu
            180                 185                 190
Gly Glu Met Leu Arg Ala Thr Ile Ala Asp Arg Tyr Tyr Arg Glu Thr
        195                 200                 205
Ala Arg Leu Ala Arg Val Leu Phe Leu His Leu Tyr Leu Phe Leu Thr
    210                 215                 220
Arg Glu Ile Leu Trp Ala Ala Tyr Ala Glu Gln Met Met Arg Pro Asp
225                 230                 235                 240
Leu Phe Asp Cys Leu Cys Cys Asp Leu Glu Ser Trp Arg Gln Leu Ala
                245                 250                 255
Gly Leu Phe Gln Pro Phe Met Phe Val Asn Gly Ala Leu Thr Val Arg
            260                 265                 270
Gly Val Pro Ile Glu Ala Arg Arg Leu Arg Glu Leu Asn His Ile Arg
        275                 280                 285
Glu His Leu Asn Leu Pro Leu Val Arg Ser Ala Ala Thr Glu Glu Pro
    290                 295                 300
Gly Ala Pro Leu Thr Thr Pro Pro Thr Leu His Gly Asn Gln Ala Arg
```

305 310 315 320
Ala Ser Gly Tyr Phe Met Val Leu Ile Arg Ala Lys Leu Asp Ser Tyr
325 330 335
Ser Ser Phe Thr Thr Ser Pro Ser Glu Ala Val Met Arg Glu His Ala
340 345 350
Tyr Ser Arg Ala Arg Thr Lys Asn Asn Tyr Gly Ser Thr Ile Glu Gly
355 360 365
Leu Leu Asp Leu Pro Asp Asp Asp Ala Pro Glu Glu Ala Gly Leu Ala
370 375 380
Ala Pro Arg Leu Ser Phe Leu Pro Ala Gly His Thr Arg Arg Leu Ser
385 390 395 400
Thr Ala Pro Pro Thr Asp Val Ser Leu Gly Asp Glu Leu His Leu Asp
405 410 415
Gly Glu Asp Val Ala Met Ala His Ala Asp Ala Leu Asp Asp Phe Asp
420 425 430
Leu Asp Met Leu Gly Asp Gly Asp Ser Pro Gly Pro Gly Phe Thr Pro
435 440 445
His Asp Ser Ala Pro Tyr Gly Ala Leu Asp Met Ala Asp Phe Glu Phe
450 455 460
Glu Gln Met Phe Thr Asp Ala Leu Gly Ile Asp Glu Tyr Gly Gly
465 470 475

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 490 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Met Asp Leu Leu Val Asp Asp Leu Phe Ala Asp Arg Asp Gly Val Ser
1 5 10 15
Pro Pro Pro Pro Arg Pro Ala Gly Gly Pro Lys Asn Thr Pro Ala Ala
20 25 30
Pro Pro Leu Tyr Ala Thr Gly Arg Leu Ser Gln Ala Gln Leu Met Pro
35 40 45
Ser Pro Pro Met Pro Val Pro Pro Ala Ala Leu Phe Asn Arg Leu Leu
50 55 60
Asp Asp Leu Gly Phe Ser Ala Gly Pro Ala Leu Cys Thr Met Leu Asp
65 70 75 80
Thr Trp Asn Glu Asp Leu Phe Ser Gly Phe Pro Thr Asn Ala Asp Met
85 90 95
Tyr Arg Glu Cys Lys Phe Leu Ser Thr Leu Pro Ser Asp Val Ile Asp
100 105 110
Trp Gly Asp Ala His Val Pro Glu Arg Ser Pro Ile Asp Ile Arg Ala
115 120 125
His Gly Asp Val Ala Phe Pro Thr Leu Pro Ala Thr Arg Asp Glu Leu
130 135 140
Pro Ser Tyr Tyr Glu Ala Met Ala Gln Phe Phe Arg Gly Glu Leu Arg
145 150 155 160
Ala Arg Glu Glu Ser Tyr Arg Thr Val Leu Ala Asn Phe Cys Ser Ala
165 170 175
Leu Tyr Arg Tyr Leu Arg Ala Ser Val Arg Gln Leu His Arg Gln Ala
180 185 190
His Met Arg Gly Arg Asn Arg Asp Leu Arg Glu Met Leu Arg Thr Thr
195 200 205
Ile Ala Asp Arg Tyr Tyr Arg Glu Thr Ala Arg Leu Ala Arg Val Leu
210 215 220
Phe Leu His Leu Tyr Leu Phe Leu Ser Arg Glu Ile Leu Trp Ala Ala
225 230 235 240
Tyr Ala Glu Gln Met Met Arg Pro Asp Leu Phe Asp Gly Leu Cys Cys
245 250 255

Asp Leu Glu Ser Trp Arg Gln Leu Ala Cys Leu Phe Gln Pro Leu Met
260 265 270
Phe Ile Asn Gly Ser Leu Thr Val Arg Gly Val Pro Val Glu Ala Arg
275 280 285
Arg Leu Arg Glu Leu Asn His Ile Arg Glu His Leu Asn Leu Pro Leu
290 295 300
Val Arg Ser Ala Ala Ala Glu Glu Pro Gly Ala Pro Leu Thr Thr Pro
305 310 315 320
Pro Val Leu Gln Gly Asn Gln Ala Arg Ser Ser Gly Tyr Phe Met Leu
325 330 335
Leu Ile Arg Ala Lys Leu Asp Ser Tyr Ser Ser Val Ala Thr Ser Glu
340 345 350
Gly Glu Ser Val Met Arg Glu His Ala Tyr Ser Arg Gly Arg Thr Arg
355 360 365
Asn Asn Tyr Gly Ser Thr Ile Glu Gly Leu Leu Asp Leu Pro Asp Asp
370 375 380
Asp Asp Ala Pro Ala Glu Ala Gly Leu Val Ala Pro Arg Met Ser Phe
385 390 395 400
Leu Ser Ala Gly Gln Arg Pro Arg Arg Leu Ser Thr Thr Ala Pro Ile
405 410 415
Thr Asp Val Ser Leu Gly Asp Glu Leu Arg Leu Asp Gly Glu Glu Val
420 425 430
Asp Met Thr Pro Ala Asp Ala Leu Asp Asp Phe Asp Leu Glu Met Leu
435 440 445
Gly Asp Val Glu Ser Pro Ser Pro Gly Met Thr His Asp Pro Val Ser
450 455 460
Tyr Gly Ala Leu Asp Val Asp Asp Phe Glu Phe Glu Gln Met Phe Thr
465 470 475 480
Asp Ala Met Gly Ile Asp Asp Phe Gly Gly
485 490

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 504 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Met Ser Gly Arg Ile Lys Thr Ala Gly Arg Ala Leu Ala Ser Gln Cys
1 5 10 15
Gly Gly Ala Ala Ala Ala Thr Met Asp Pro Tyr Asp Ala Ile Glu Ala
20 25 30
Phe Asp Asp Ser Leu Leu Gly Ser Pro Leu Ala Ala Gly Pro Leu Tyr
35 40 45
Asp Gly Pro Ser Pro Ala Arg Phe Ala Leu Pro Pro Pro Arg Pro Ala
50 55 60
Pro Leu Ala Ala Leu Leu Glu Arg Met Gln Ala Glu Leu Gly Phe Pro
65 70 75 80
Asp Gly Pro Ala Leu Leu Arg Ala Met Glu Arg Trp Asn Glu Asp Leu
85 90 95
Phe Ser Cys Leu Pro Thr Asn Ala Asp Leu Tyr Ala Asp Ala Ala Leu
100 105 110
Leu Ser Ala Asp Ala Asp Ala Val Val Gly Ala Met Tyr Leu Ala Val
115 120 125
Pro Gly Asp Ala Glu Arg Leu Asp Leu Asn Ala His Ala Asn Gln Pro
130 135 140
Leu Pro Ala Pro Pro Ala Ser Glu Glu Gly Leu Pro Glu Tyr Val Ala
145 150 155 160
Gly Val Gln Ala His Phe Leu Ala Glu Leu Arg Ala Arg Glu Glu Arg
165 170 175
Tyr Ala Gly Leu Phe Leu Gly Tyr Cys Arg Ala Leu Leu Gln His Leu

```
                180                     185                     190
Arg Ala Thr Ala Ala Arg Gly Arg Gly Ala Ala Gly Ala Gly Ala Gln
            195                 200                     205
Ala Asp Arg Leu Arg Gln Leu Val Ala Ala Arg Tyr Tyr Arg Glu Ala
    210                 215                     220
Ser Arg Leu Ala Arg Leu Ala Phe Ala His Met Tyr Val Ala Thr Ala
225                 230                     235                 240
Arg Glu Val Ser Trp Arg Leu His Ser Gln Gln Ser Gln Ala Gln Gly
                245                     250                 255
Val Phe Val Ser Leu Tyr Tyr Ala Trp Pro Gln Arg Arg Gln Phe Thr
            260                     265                 270
Cys Leu Phe His Pro Val Leu Phe Asn His Gly Val Val Ala Leu Glu
        275                 280                     285
Asp Gly Phe Leu Asp Ala Ala Glu Leu Arg Arg Leu Asn Tyr Arg Arg
    290                 295                     300
Arg Glu Leu Gly Leu Pro Leu Val Arg Ala Gly Leu Val Glu Val Glu
305                 310                     315                 320
Val Gly Pro Leu Val Glu Glu Pro Pro Phe Ser Gly Ser Leu Pro Arg
                325                     330                 335
Ala Leu Gly Phe Leu Asn Tyr Gln Val Arg Ala Lys Met Gly Ala Pro
            340                     345                 350
Ala Glu Ala Gly Gly Gly Trp Arg Arg Ser Gly Ser Thr Arg Thr Arg
        355                 360                     365
Gly Arg Ala Ala Arg Ser Thr Thr Gly Arg Leu Gln Arg Pro Cys Cys
    370                 375                     380
Gly Pro Arg Arg Arg Ala Lys Cys Cys Arg Ala Thr Pro Arg Gln Arg
385                 390                     395                 400
Leu Arg Ala Arg Gly Glu Pro Arg His Thr Ser Gly Ser Gly Ala Phe
                405                     410                 415
Ser Gln Gly Arg Arg Pro Gly Arg Val Cys Arg Leu Gly Trp Ala Cys
            420                     425                 430
Lys Ala Arg Ser Gly Pro Ala Arg Gly Gly Pro Gly Pro Ser Pro Val
        435                 440                     445
Arg Ser Gly Leu Gly Leu Ser Arg Ala Arg Gly Ser Pro Gly Pro Gly
    450                 455                     460
Pro Ala Cys Gly Gly Pro Ser Arg Ala Arg Gly Gly Arg Arg Arg Ala
465                 470                     475                 480
Ser Pro Ala Asn Pro Phe Gly Gly Thr Tyr Asp Ala Leu Leu Gly Asp
                485                     490                     495
Arg Leu Asn Gln Leu Leu Asp Phe
            500
```

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 410 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

```
Met Glu Cys Asn Leu Gly Thr Glu His Pro Ser Thr Asp Thr Trp Asn
 1               5                  10                  15
Arg Ser Lys Thr Glu Gln Ala Val Val Asp Ala Phe Asp Glu Ser Leu
            20                  25                  30
Phe Gly Asp Val Ala Ser Asp Ile Gly Phe Glu Thr Ser Leu Tyr Ser
        35                  40                  45
His Ala Val Lys Thr Ala Pro Ser Pro Pro Trp Val Ala Ser Pro Lys
    50                  55                  60
Ile Leu Tyr Gln Gln Leu Ile Arg Asp Leu Asp Phe Ser Glu Gly Pro
65                  70                  75                  80
Arg Leu Leu Ser Cys Leu Glu Thr Trp Asn Glu Asp Leu Phe Ser Cys
                85                  90                  95
```

```
Phe Pro Ile Asn Glu Asp Leu Tyr Ser Asp Met Met Val Leu Ser Pro
            100                 105                 110
Asp Pro Asp Asp Val Ile Ser Thr Val Ser Thr Lys Asp His Val Glu
        115                 120                 125
Met Phe Asn Leu Thr Thr Arg Gly Ser Val Arg Leu Pro Ser Pro Pro
    130                 135                 140
Lys Gln Pro Thr Gly Leu Pro Ala Tyr Val Gln Glu Val Gln Asp Ser
145                 150                 155                 160
Phe Thr Val Glu Leu Arg Ala Arg Glu Glu Ala Tyr Thr Lys Leu Leu
                165                 170                 175
Val Thr Tyr Cys Lys Ser Ile Ile Arg Tyr Leu Gln Gly Thr Ala Lys
            180                 185                 190
Arg Thr Thr Ile Gly Leu Asn Ile Gln Asn Pro Asp Gln Lys Ala Tyr
        195                 200                 205
Thr Gln Leu Arg Gln Ser Ile Leu Leu Arg Tyr Tyr Arg Glu Val Ala
    210                 215                 220
Ser Leu Ala Arg Leu Leu Tyr Leu His Leu Tyr Leu Thr Val Thr Arg
225                 230                 235                 240
Glu Phe Ser Trp Arg Leu Tyr Ala Ser Gln Ser Ala His Pro Asp Val
                245                 250                 255
Phe Ala Ala Leu Lys Phe Thr Trp Thr Glu Arg Arg Gln Phe Thr Cys
            260                 265                 270
Ala Phe His Pro Val Leu Cys Asn His Gly Ile Val Leu Leu Glu Gly
        275                 280                 285
Lys Pro Leu Thr Ala Ser Ala Leu Arg Glu Ile Asn Tyr Arg Arg Arg
    290                 295                 300
Glu Leu Gly Leu Pro Leu Val Arg Cys Gly Leu Val Glu Glu Asn Lys
305                 310                 315                 320
Ser Pro Leu Val Gln Gln Pro Ser Phe Ser Val His Leu Pro Arg Ser
                325                 330                 335
Val Gly Phe Leu Thr His His Ile Lys Arg Lys Leu Asp Ala Tyr Ala
            340                 345                 350
Val Lys His Pro Gln Glu Pro Arg His Val Arg Ala Asp His Pro Tyr
        355                 360                 365
Ala Lys Val Val Glu Asn Arg Asn Tyr Gly Ser Ser Ile Glu Ala Met
    370                 375                 380
Ile Leu Ala Pro Pro Ser Pro Ser Glu Ile Leu Pro Gly Asp Pro Pro
385                 390                 395                 400
Arg Pro Pro Thr Cys Gly Phe Leu Thr Arg
                405                 410
```

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 454 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

```
Met Ala Ala Asn Ile Ala Met Phe Ala Asp Ile Glu Asp Tyr Asp Asp
 1               5                  10                  15
Thr Arg Ser Cys Glu Tyr Gly Tyr Gly Thr Cys Glu Leu Met Asp Val
            20                  25                  30
Asp Gly Val Val Ala Ser Phe Asp Glu Gly Met Leu Ser Ala Ser Glu
        35                  40                  45
Ser Ile Tyr Ser Ser Pro Ala Gln Lys Arg Leu Ala Leu Pro Pro Pro
    50                  55                  60
Lys Ala Thr Ser Pro Thr Ala Leu Tyr Gln Arg Leu Gln Ala Glu Leu
65                  70                  75                  80
Gly Phe Pro Glu Gly Gln Ala Met Leu Phe Ala Met Glu Lys Trp Asn
                85                  90                  95
Glu Asp Met Phe Ser Ala Ile Pro Val His Val Asp Leu Tyr Thr Glu
```

```
                100                           105                           110
Ile Ala Leu Leu Ser Thr Ser Val Asn Glu Val Val Lys Ala Gly Leu
            115                           120                           125
Asp Ser Leu Pro Ile Pro Thr Asn Tyr Ile Pro Glu Val Asp Leu Asn
    130                           135                           140
Ala His Gly Ser Glu Pro Phe Pro Glu Val Pro Ala Leu Glu Asp Glu
145                           150                           155                           160
Leu Glu Thr Tyr Val Ile Ser Ala Gln Arg Phe Tyr Leu Ser Glu Leu
                        165                           170                           175
Arg Ala Arg Glu Glu His Tyr Ser Arg Leu Leu Arg Gly Tyr Cys Val
                180                           185                           190
Ala Leu Leu His Tyr Leu Tyr Gly Ser Ala Lys Arg Gln Leu Arg Gly
            195                           200                           205
Ala Gly Ser Asp Ser Ala Leu Met His Lys Phe Lys Gln Val Val Arg
    210                           215                           220
Asp Arg Tyr Tyr Arg Glu Thr Ala Asn Leu Ala Arg Leu Leu Tyr Leu
225                           230                           235                           240
His Leu Tyr Ile Ser Val Thr Arg Glu Val Ser Trp Arg Leu His Ala
                        245                           250                           255
Ser Gln Val Val Asn Gln Gly Ile Phe Val Ser Leu His Tyr Thr Trp
                260                           265                           270
Pro Gln Arg Arg Lys Phe Glu Cys Leu Phe His Pro Val Leu Phe Asn
            275                           280                           285
His Gly Val Val Ile Leu Glu Asn Asp Pro Leu Glu Phe Asn Asp Leu
    290                           295                           300
Gln Arg Ile Asn Tyr Arg Arg Arg Glu Leu Gly Leu Pro Leu Ile Arg
305                           310                           315                           320
Ala Gly Leu Ile Glu Glu Glu Asn Leu Pro Leu Glu Ser Glu Pro Thr
                        325                           330                           335
Phe Ser Gly Lys Leu Pro Arg Thr Ile Gly Phe Leu Thr His Gln Ile
                340                           345                           350
Arg Thr Lys Met Glu Ala Tyr Ser Asn Ala His Pro Ser Thr Pro Leu
            355                           360                           365
Phe Pro Leu Ala Glu His Ser Tyr Ser Lys Arg Ile Asp Gly Arg Leu
    370                           375                           380
Ser Tyr Gly Thr Thr Ala Glu Ala Met Met Asp Pro Pro Ser Pro Ser
385                           390                           395                           400
Ala Val Leu Pro Gly Asp Pro Val Pro Pro Leu Thr Val Gly Ile Arg
                        405                           410                           415
Gln Thr Ala Glu Thr Leu Ala Leu Pro Ser Asn Leu Thr Leu Gln Ser
                420                           425                           430
Met Glu Thr Asp Val Leu Asp Tyr Ser Ser Ile Ser Gly Asp Glu Leu
            435                           440                           445
Asn Gln Met Phe Asp Ile
    450
```

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 479 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

```
Met Cys Leu Leu His Ile Ser Leu Pro Tyr Leu Ser Cys Ala Leu Leu
1                       5                             10                            15
Pro Gly Trp Tyr Phe Asp Ala Arg Pro Ala Ala Ser Ile Val Met Phe
                20                            25                            30
Ala Ala Ala Glu Glu Asn Asp Asp Pro Tyr Pro Gly Lys Ser Gly Tyr
            35                            40                            45
Asn Asp Thr Cys Glu Leu Met Asp Met Asp Gly Ala Val Ala Ser Phe
    50                            55                            60
```

Asp Glu Gly Met Leu Ser Ala Ile Glu Ser Val Tyr Ser Ile Pro Thr
65 70 75 80
Lys Lys Arg Leu Ala Leu Pro Pro Pro Lys Ala Ala Ser Pro Gly Ala
85 90 95
Leu Tyr Gln Arg Leu Gln Gly Glu Leu Gly Phe Pro Glu Gly Gln Thr
100 105 110
Leu Leu Ser Ala Met Glu Lys Trp Asn Glu Asp Met Phe Ser Ala Leu
115 120 125
Pro Gly His Val Asp Leu Tyr Thr Glu Ile Ala Leu Leu Ser Thr Ser
130 135 140
Val Asp Glu Val Val Arg Ala Gly Leu Asp Ser Leu Pro Thr Pro Ser
145 150 155 160
His Tyr Ser Pro Glu Val Asp Leu Asn Ala His Gly Asp Glu Pro Phe
165 170 175
Pro Glu Val Pro Ala Leu Glu Asp Asp Leu Glu Ile Tyr Val Ile Ser
180 185 190
Ala Gln Arg Phe Tyr Leu Ser Glu Leu Arg Thr Arg Glu Glu His Tyr
195 200 205
Ala Arg Leu Leu Arg Gly Tyr Cys Val Ala Leu Leu His Tyr Leu Tyr
210 215 220
Gly Ser Ala Lys Arg Gln Leu Arg Gly Ser Gly Ser Asp Ala Ser Leu
225 230 235 240
Met His Lys Phe Lys Gln Val Val Arg Asp Arg Tyr Tyr Arg Glu Ala
245 250 255
Ala Asn Leu Ala Arg Leu Leu Tyr Leu His Leu Tyr Val Ser Val Thr
260 265 270
Arg Glu Val Ser Trp Arg Leu His Ala Ser Gln Val Ile Asn Gln Gly
275 280 285
Val Phe Val Ser Leu His Tyr Phe Trp Ala Gln Arg Arg Lys Phe Glu
290 295 300
Cys Leu Phe His Pro Val Leu Phe Asn His Gly Val Val Ile Leu Glu
305 310 315 320
Asn Asp Pro Leu Glu Phe His Asp Leu Gln Arg Ile Asn Tyr Arg Arg
325 330 335
Arg Glu Leu Gly Leu Pro Leu Ile Arg Ala Gly Leu Ile Glu Glu Glu
340 345 350
Asn Ser Pro Leu Glu Ala Glu Pro Leu Phe Ser Gly Lys Leu Pro Arg
355 360 365
Thr Ile Gly Phe Leu Thr His Gln Ile Arg Thr Lys Met Glu Ala Tyr
370 375 380
Ser Asp Ala His Pro Ala Thr Pro Leu Phe Pro Leu Ala Glu His Ser
385 390 395 400
Tyr Ser Lys Arg Ile Gly Gly Arg Leu Ser Tyr Gly Thr Thr Thr Glu
405 410 415
Ala Met Met Asp Pro Pro Ser Pro Ser Ala Val Leu Pro Gly Asp Pro
420 425 430
Val Pro Pro Leu Thr Val Gly Val Arg Gln Thr Ala Ala Thr Leu Ala
435 440 445
Ile Pro Ser Asn Leu Thr Leu Gln Ser Met Glu Thr Asp Gly Leu Asp
450 455 460
Tyr Ser Ser Met Thr Gly Asp Glu Leu Asn Gln Met Phe Asp Ile
465 470 475

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 752 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

Met Gly Pro Leu Met Val Leu Phe Cys Leu Leu Phe Leu Tyr Pro Gly

1 5 10 15
Leu Ala Asp Ser Ala Pro Ser Cys Pro Gln Asn Val Asn Ile Ser Gly
20 25 30
Gly Thr Phe Thr Leu Ser His Gly Trp Ala Pro Gly Ser Leu Leu Thr
35 40 45
Tyr Ser Cys Pro Gln Gly Leu Tyr Pro Ser Pro Ala Ser Arg Leu Cys
50 55 60
Lys Ser Ser Gly Gln Trp Gln Thr Pro Gly Ala Thr Arg Ser Leu Ser
65 70 75 80
Lys Ala Val Cys Lys Pro Val Arg Cys Pro Ala Pro Val Ser Phe Glu
85 90 95
Asn Gly Ile Tyr Thr Pro Arg Leu Gly Ser Tyr Pro Val Gly Gly Asn
100 105 110
Val Ser Phe Glu Cys Glu Asp Gly Phe Ile Leu Arg Gly Ser Pro Val
115 120 125
Arg Gln Cys Arg Pro Asn Gly Met Trp Asp Gly Glu Thr Ala Val Cys
130 135 140
Asp Asn Gly Ala Gly His Cys Pro Asn Pro Gly Ile Ser Leu Gly Ala
145 150 155 160
Val Arg Thr Gly Phe Arg Phe Gly His Gly Asp Lys Val Arg Tyr Arg
165 170 175
Cys Ser Ser Asn Leu Val Leu Thr Gly Ser Ser Glu Arg Glu Cys Gln
180 185 190
Gly Asn Gly Val Trp Ser Gly Thr Glu Pro Ile Cys Arg Gln Pro Tyr
195 200 205
Ser Tyr Asp Phe Pro Glu Asp Val Ala Pro Ala Leu Gly Thr Ser Phe
210 215 220
Ser His Met Leu Gly Ala Thr Asn Pro Thr Gln Lys Thr Lys Glu Ser
225 230 235 240
Leu Gly Arg Lys Ile Gln Ile Gln Arg Ser Gly His Leu Asn Leu Tyr
245 250 255
Leu Leu Leu Asp Cys Ser Gln Ser Val Ser Glu Asn Asp Phe Leu Ile
260 265 270
Phe Lys Glu Ser Ala Ser Leu Met Val Asp Arg Ile Phe Ser Phe Glu
275 280 285
Ile Asn Val Ser Val Ala Ile Ile Thr Phe Ala Ser Glu Pro Lys Val
290 295 300
Leu Met Ser Val Leu Asn Asp Asn Ser Arg Asp Met Thr Glu Val Ile
305 310 315 320
Ser Ser Leu Glu Asn Ala Asn Tyr Lys Asp His Glu Asn Gly Thr Gly
325 330 335
Thr Asn Thr Tyr Ala Ala Leu Asn Ser Val Tyr Leu Met Met Asn Asn
340 345 350
Gln Met Arg Leu Leu Gly Met Glu Thr Met Ala Trp Gln Glu Ile Arg
355 360 365
His Ala Ile Ile Leu Leu Thr Asp Gly Lys Ser Asn Met Gly Gly Ser
370 375 380
Pro Lys Thr Ala Val Asp His Ile Arg Glu Ile Leu Asn Ile Asn Gln
385 390 395 400
Lys Arg Asn Asp Tyr Leu Asp Ile Tyr Ala Ile Gly Val Gly Lys Leu
405 410 415
Asp Val Asp Trp Arg Glu Leu Asn Glu Leu Gly Ser Lys Lys Asp Gly
420 425 430
Glu Arg His Ala Phe Ile Leu Gln Asp Thr Lys Ala Leu His Gln Val
435 440 445
Phe Glu His Met Leu Asp Val Ser Lys Leu Thr Asp Thr Ile Cys Gly
450 455 460
Val Gly Asn Met Ser Ala Asn Ala Ser Asp Gln Glu Arg Thr Pro Trp
465 470 475 480
His Val Thr Ile Lys Pro Lys Ser Gln Glu Thr Cys Arg Gly Ala Leu
485 490 495
Ile Ser Asp Gln Trp Val Leu Thr Ala Ala His Cys Phe Arg Asp Gly
500 505 510
Asn Asp His Ser Leu Trp Arg Val Asn Val Gly Asp Pro Lys Ser Gln
515 520 525
Trp Gly Lys Glu Leu Leu Ile Glu Lys Ala Val Ile Ser Pro Gly Phe
530 535 540
Asp Val Phe Ala Lys Lys Asn Gln Gly Ile Leu Glu Phe Tyr Gly Asp
545 550 555 560
Asp Ile Ala Leu Leu Lys Leu Ala Gln Lys Val Lys Met Ser Thr His

```
                565                     570                     575
Ala Arg Pro Ile Cys Leu Pro Cys Thr Met Glu Ala Asn Leu Ala Leu
            580                     585                     590
Arg Arg Pro Gln Gly Ser Thr Cys Arg Asp His Glu Asn Glu Leu Leu
        595                     600                     605
Asn Lys Gln Ser Val Pro Ala His Phe Val Ala Leu Asn Gly Ser Lys
    610                     615                     620
Leu Asn Ile Asn Leu Lys Met Gly Val Glu Trp Thr Ser Cys Ala Glu
625                     630                     635                     640
Val Val Ser Gln Glu Lys Thr Met Phe Pro Asn Leu Thr Asp Val Arg
                645                     650                     655
Glu Val Val Thr Asp Gln Phe Leu Cys Ser Gly Thr Gln Glu Asp Glu
            660                     665                     670
Ser Pro Cys Lys Gly Glu Ser Gly Gly Ala Val Phe Leu Glu Arg Arg
        675                     680                     685
Phe Arg Phe Phe Gln Val Gly Leu Val Ser Trp Gly Leu Tyr Asn Pro
    690                     695                     700
Cys Leu Gly Ser Ala Asp Lys Asn Ser Arg Lys Arg Ala Pro Arg Ser
705                     710                     715                     720
Lys Val Pro Pro Pro Arg Asp Phe His Ile Asn Leu Phe Arg Met Gln
                725                     730                     735
Pro Trp Leu Arg Gln His Leu Gly Asp Val Leu Asn Phe Leu Pro Leu
            740                     745                     750
```

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 760 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
Met Ala Pro Leu Leu Ala Leu Phe Tyr Leu Leu Gln Leu Gly Pro Gly
 1               5                  10                  15
Leu Ala Ala Leu Phe Cys Asn Gln Asn Val Asn Ile Thr Gly Gly Asn
            20                  25                  30
Phe Thr Leu Ser His Gly Trp Ala Pro Gly Ser Leu Leu Ile Tyr Ser
        35                  40                  45
Cys Pro Leu Gly Arg Tyr Pro Ser Pro Ala Trp Arg Lys Cys Gln Ser
    50                  55                  60
Asn Gly Gln Trp Leu Thr Pro Arg Ser Ser Ser His His Thr Leu Arg
65                  70                  75                  80
Ser Ser Arg Met Val Lys Ala Val Cys Lys Pro Val Arg Cys Leu Ala
                85                  90                  95
Pro Ser Ser Phe Glu Asn Gly Ile Tyr Phe Pro Arg Leu Val Ser Tyr
            100                 105                 110
Pro Val Gly Ser Asn Val Ser Phe Glu Cys Asp Glu Asp Phe Thr Leu
        115                 120                 125
Arg Gly Ser Pro Val Arg Tyr Cys Arg Pro Asn Gly Leu Trp Asp Gly
    130                 135                 140
Glu Thr Ala Val Cys Asp Asn Gly Ala Ser His Cys Pro Asn Pro Gly
145                 150                 155                 160
Ile Ser Val Gly Thr Ala Arg Thr Gly Leu Asn Phe Asp Leu Gly Asp
                165                 170                 175
Lys Val Arg Tyr Arg Cys Ser Ser Ser Asn Met Val Leu Thr Gly Ser
            180                 185                 190
Ala Glu Arg Glu Cys Gln Ser Asn Gly Val Tr  Ser Gly Ser Glu Pro
        195                 200                 205
Ile Cys Arg Gln Pro Tyr Ser Tyr Asp Phe Pro Glu Asp Val Ala Ser
    210                 215                 220
Ala Leu Asp Thr Ser Leu Thr Asn Leu Leu Gly Ala Thr Asn Pro Thr
225                 230                 235                 240
```

```
Gln Asn Leu Leu Thr Lys Ser Leu Gly Arg Lys Ile Ile Ile Gln Arg
                245                 250                 255
Ser Gly His Leu Asn Leu Tyr Leu Leu Leu Asp Ala Ser Gln Ser Val
            260                 265                 270
Thr Glu Lys Asp Phe Asp Ile Phe Lys Lys Ser Ala Glu Leu Met Val
        275                 280                 285
Glu Arg Ile Phe Ser Phe Glu Val Asn Val Thr Val Ala Ile Ile Thr
    290                 295                 300
Phe Ala Ser Gln Pro Lys Thr Ile Met Ser Ile Leu Ser Glu Arg Ser
305                 310                 315                 320
Gln Asp Val Thr Glu Val Ile Thr Ser Leu Asp Ser Ala Ser Tyr Lys
                325                 330                 335
Asp His Glu Asn Ala Thr Gly Ala Asn Thr Tyr Glu Val Leu Ile Arg
            340                 345                 350
Val Tyr Ser Met Met Gln Thr Gln Met Asp Arg Leu Gly Met Glu Thr
        355                 360                 365
Ser Ala Trp Lys Glu Ile Arg His Thr Ile Ile Leu Leu Thr Asp Gly
    370                 375                 380
Lys Ser Asn Met Gly Asp Ser Pro Lys Lys Ala Val Thr Arg Ile Arg
385                 390                 395                 400
Glu Leu Leu Ser Ile Glu Gln Asn Arg Asp Asp Tyr Leu Asp Ile Tyr
                405                 410                 415
Ala Ile Gly Val Gly Lys Leu Asp Val Asp Trp Lys Glu Leu Asn Glu
            420                 425                 430
Leu Gly Ser Lys Lys Asp Gly Glu Arg His Ala Phe Ile Leu Gln Asp
        435                 440                 445
Ala Lys Ala Leu Gln Gln Ile Phe Glu His Met Leu Asp Val Ser Lys
    450                 455                 460
Leu Thr Asp Thr Ile Cys Gly Val Gly Asn Met Ser Ala Asn Ala Ser
465                 470                 475                 480
Asp Gln Glu Arg Thr Pro Trp Gln Val Thr Phe Lys Pro Lys Ser Lys
                485                 490                 495
Glu Thr Cys Gln Gly Ser Leu Ile Ser Asp Gln Trp Val Leu Thr Ala
            500                 505                 510
Ala His Cys Phe His Asp Ile Gln Met Glu Asp His His Leu Trp Arg
        515                 520                 525
Val Asn Val Gly Asp Pro Thr Ser Gln His Gly Lys Glu Phe Leu Val
    530                 535                 540
Glu Asp Val Ile Ile Ala Pro Gly Phe Asn Val His Ala Lys Arg Lys
545                 550                 555                 560
Gln Gly Ile Ser Glu Phe Tyr Ala Asp Asp Ile Ala Leu Leu Lys Leu
                565                 570                 575
Ser Arg Lys Val Lys Met Ser Thr His Ala Arg Pro Ile Cys Leu Pro
            580                 585                 590
Cys Thr Val Gly Ala Asn Met Ala Leu Arg Arg Ser Pro Gly Ser Thr
        595                 600                 605
Cys Lys Asp His Glu Thr Glu Leu Leu Ser Gln Gln Lys Val Pro Ala
    610                 615                 620
His Phe Val Ala Leu Asn Gly Asn Arg Leu Asn Ile Asn Leu Arg Thr
625                 630                 635                 640
Gly Pro Glu Trp Thr Arg Cys Ile Gln Ala Val Ser Gln Asn Lys Asn
                645                 650                 655
Ile Phe Pro Ser Leu Thr Asn Val Ser Glu Val Val Thr Asp Gln Phe
            660                 665                 670
Leu Cys Ser Gly Met Glu Glu Glu Asp Asp Asn Pro Cys Lys Gly Glu
        675                 680                 685
Ser Gly Gly Ala Val Phe Leu Gly Arg Arg Tyr Arg Phe Phe Gln Val
    690                 695                 700
Gly Leu Val Ser Trp Gly Leu Phe Asp Pro Cys His Gly Ser Ser Asn
705                 710                 715                 720
Lys Asn Leu Arg Lys Lys Pro Pro Arg Gly Val Leu Pro Arg Asp Phe
                725                 730                 735
His Ile Ser Leu Phe Arg Leu Gln Pro Trp Leu Arg Gln His Leu Asp
            740                 745                 750
Gly Val Leu Asp Phe Leu Pro Leu
        755                 760
```

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 764 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

Met Gly Ser Asn Leu Ser Pro Gln Leu Cys Leu Met Pro Phe Ile Leu
1 5 10 15
Gly Leu Leu Ser Gly Gly Val Thr Thr Thr Pro Trp Ser Leu Ala Arg
20 25 30
Pro Gln Gly Ser Cys Ser Leu Glu Gly Val Glu Ile Lys Gly Gly Ser
35 40 45
Phe Arg Leu Leu Gln Glu Gly Gln Ala Leu Glu Tyr Val Cys Pro Ser
50 55 60
Gly Phe Tyr Pro Tyr Pro Val Gln Thr Arg Thr Cys Arg Ser Thr Gly
65 70 75 80
Ser Trp Ser Thr Leu Lys Thr Gln Asp Gln Lys Thr Val Arg Lys Ala
85 90 95
Glu Cys Arg Ala Ile His Cys Pro Arg Pro His Asp Phe Glu Asn Gly
100 105 110
Glu Tyr Trp Pro Arg Ser Pro Tyr Tyr Asn Val Ser Asp Glu Ile Ser
115 120 125
Phe His Cys Tyr Asp Gly Tyr Thr Leu Arg Gly Ser Ala Asn Arg Thr
130 135 140
Cys Gln Val Asn Gly Arg Trp Ser Gly Gln Thr Ala Ile Cys Asp Asn
145 150 155 160
Gly Ala Gly Tyr Cys Ser Asn Pro Gly Ile Pro Ile Gly Thr Arg Lys
165 170 175
Val Gly Ser Gln Tyr Arg Leu Glu Asp Ser Val Thr Tyr His Cys Ser
180 185 190
Arg Gly Leu Thr Leu Arg Gly Ser Gln Arg Arg Thr Cys Gln Glu Gly
195 200 205
Gly Ser Trp Ser Gly Thr Glu Pro Ser Cys Gln Asp Ser Phe Met Tyr
210 215 220
Asp Thr Pro Gln Glu Val Ala Glu Ala Phe Leu Ser Ser Leu Thr Glu
225 230 235 240
Thr Ile Glu Gly Val Asp Ala Glu Asp Gly His Gly Pro Gly Glu Gln
245 250 255
Gln Lys Arg Lys Ile Val Leu Asp Pro Ser Gly Ser Met Asn Ile Tyr
260 265 270
Leu Val Leu Asp Gly Ser Asp Ser Ile Gly Ala Ser Asn Phe Thr Gly
275 280 285
Ala Lys Lys Cys Leu Val Asn Leu Ile Glu Lys Val Ala Ser Tyr Gly
290 295 300
Val Lys Pro Arg Tyr Gly Leu Val Thr Tyr Ala Thr Tyr Pro Lys Ile
305 310 315 320
Trp Val Lys Val Ser Glu Ala Asp Ser Ser Asn Ala Asp Trp Val Thr
325 330 335
Lys Gln Leu Asn Glu Ile Asn Tyr Glu Asp His Lys Leu Lys Ser Gly
340 345 350
Thr Asn Thr Lys Lys Ala Leu Gln Ala Val Tyr Ser Met Met Ser Trp
355 360 365
Pro Asp Asp Val Pro Pro Glu Gly Trp Asn Arg Thr Arg His Val Ile
370 375 380
Ile Leu Met Thr Asp Gly Leu His Asn Met Gly Gly Asp Pro Ile Thr
385 390 395 400
Val Ile Asp Glu Ile Arg Asp Leu Leu Tyr Ile Gly Lys Asp Arg Lys
405 410 415
Asn Pro Arg Glu Asp Tyr Leu Asp Val Tyr Val Phe Gly Val Gly Pro
420 425 430
Leu Val Asn Gln Val Asn Ile Asn Ala Leu Ala Ser Lys Lys Asp Asn
435 440 445
Glu Gln His Val Phe Lys Val Lys Asp Met Glu Asn Leu Glu Asp Val

```
    450                 455                 460
Phe Tyr Gln Met Ile Asp Glu Ser Gln Ser Leu Ser Leu Cys Gly Met
465                 470                 475                 480
Val Trp Glu His Arg Lys Gly Thr Asp Tyr His Lys Gln Pro Trp Gln
                485                 490                 495
Ala Lys Ile Ser Val Ile Arg Pro Ser Lys Gly His Glu Ser Cys Met
            500                 505                 510
Gly Ala Val Val Ser Glu Tyr Phe Val Leu Thr Ala Ala His Cys Phe
        515                 520                 525
Thr Val Asp Asp Lys Glu His Ser Ile Lys Val Ser Val Gly Gly Glu
    530                 535                 540
Lys Arg Asp Leu Glu Ile Glu Val Val Leu Phe His Pro Asn Tyr Asn
545                 550                 555                 560
Ile Asn Gly Lys Lys Glu Ala Gly Ile Pro Glu Phe Tyr Asp Tyr Asp
                565                 570                 575
Val Ala Leu Ile Lys Leu Lys Asn Lys Leu Lys Tyr Gly Gln Thr Ile
            580                 585                 590
Arg Pro Ile Cys Leu Pro Cys Thr Glu Gly Thr Thr Arg Ala Leu Arg
        595                 600                 605
Leu Pro Pro Thr Thr Thr Cys Gln Gln Gln Lys Glu Glu Leu Leu Pro
    610                 615                 620
Ala Gln Asp Ile Lys Ala Leu Phe Val Ser Glu Glu Glu Lys Lys Leu
625                 630                 635                 640
Thr Arg Lys Glu Val Tyr Ile Lys Asn Gly Asp Lys Lys Gly Ser Cys
                645                 650                 655
Glu Arg Asp Ala Gln Tyr Ala Pro Gly Tyr Asp Lys Val Lys Asp Ile
            660                 665                 670
Ser Glu Val Val Thr Pro Arg Phe Leu Cys Thr Gly Gly Val Ser Pro
        675                 680                 685
Tyr Ala Asp Pro Asn Thr Cys Arg Gly Asp Ser Gly Gly Pro Leu Ile
    690                 695                 700
Val His Lys Arg Ser Arg Phe Ile Gln Val Gly Val Ile Ser Trp Gly
705                 710                 715                 720
Val Val Asp Val Cys Lys Asn Gln Lys Arg Gln Lys Gln Val Pro Ala
                725                 730                 735
His Ala Arg Asp Phe His Ile Asn Leu Phe Gln Val Leu Pro Trp Leu
            740                 745                 750
Lys Glu Lys Leu Gln Asp Glu Asp Leu Gly Phe Leu
        755                 760
```

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 761 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

```
Met Glu Ser Pro Gln Leu Cys Leu Val Leu Leu Val Leu Gly Phe Ser
1               5                   10                  15
Ser Gly Gly Val Ser Ala Thr Pro Val Leu Glu Ala Arg Pro Gln Val
            20                  25                  30
Ser Cys Ser Leu Glu Gly Val Glu Ile Lys Gly Gly Ser Phe Gln Leu
        35                  40                  45
Leu Gln Gly Gly Gln Ala Leu Glu Tyr Leu Cys Pro Ser Gly Phe Tyr
    50                  55                  60
Pro Tyr Pro Val Gln Thr Arg Thr Cys Arg Ser Thr Gly Ser Trp Ser
65                  70                  75                  80
Asp Leu Gln Thr Arg Asp Gln Lys Ile Val Gln Lys Ala Glu Cys Arg
                85                  90                  95
Ala Ile Arg Cys Pro Arg Pro Gln Asp Phe Glu Asn Gly Glu Phe Trp
            100                 105                 110
```

Pro Arg Ser Pro Phe Tyr Asn Leu Ser Asp Gln Ile Ser Phe Gln Cys
115 120 125
Tyr Asp Gly Tyr Val Leu Arg Gly Ser Ala Asn Arg Thr Cys Gln Glu
130 135 140
Asn Gly Arg Trp Asp Gly Gln Thr Ala Ile Cys Asp Asp Gly Ala Gly
145 150 155 160
Tyr Cys Pro Asn Pro Gly Ile Pro Ile Gly Thr Arg Lys Val Gly Ser
165 170 175
Gln Tyr Arg Leu Glu Asp Ile Val Thr Tyr His Cys Ser Arg Gly Leu
180 185 190
Val Leu Arg Gly Ser Gln Lys Arg Lys Cys Gln Glu Gly Gly Ser Trp
195 200 205
Ser Gly Thr Glu Pro Ser Cys Gln Asp Ser Phe Met Tyr Asp Ser Pro
210 215 220
Gln Glu Val Ala Glu Ala Phe Leu Ser Ser Leu Thr Glu Thr Ile Glu
225 230 235 240
Gly Ala Asp Ala Glu Asp Gly His Ser Pro Gly Glu Gln Gln Lys Arg
245 250 255
Lys Ile Val Leu Asp Pro Ser Gly Ser Met Asn Ile Tyr Leu Val Leu
260 265 270
Asp Gly Ser Asp Ser Ile Gly Ser Ser Asn Phe Thr Gly Ala Lys Arg
275 280 285
Cys Leu Thr Asn Leu Ile Glu Lys Val Ala Ser Tyr Gly Val Arg Pro
290 295 300
Arg Tyr Gly Leu Leu Thr Tyr Ala Thr Val Pro Lys Val Leu Val Arg
305 310 315 320
Val Ser Asp Glu Arg Ser Ser Asp Ala Asp Trp Val Thr Glu Lys Leu
325 330 335
Asn Gln Ile Ser Tyr Glu Asp His Lys Leu Lys Ser Gly Thr Asn Thr
340 345 350
Lys Arg Ala Leu Gln Ala Val Tyr Ser Met Met Ser Trp Ala Gly Asp
355 360 365
Ala Pro Pro Glu Gly Trp Asn Arg Thr Arg His Val Ile Ile Ile Met
370 375 380
Thr Asp Gly Leu His Asn Met Gly Gly Asn Pro Val Thr Val Ile Gln
385 390 395 400
Asp Ile Arg Ala Leu Leu Asp Ile Gly Arg Asp Pro Lys Asn Pro Arg
405 410 415
Glu Asp Tyr Leu Asp Val Tyr Val Phe Gly Val Gly Pro Leu Val Asp
420 425 430
Ser Val Asn Ile Asn Ala Leu Ala Ser Lys Lys Asp Asn Glu His His
435 440 445
Val Phe Lys Val Lys Asp Met Glu Asp Leu Glu Asn Val Phe Tyr Gln
450 455 460
Met Ile Asp Glu Thr Lys Ser Leu Ser Leu Cys Gly Met Val Trp Glu
465 470 475 480
His Lys Lys Gly Asn Asp Tyr His Lys Gln Pro Trp Gln Ala Lys Ile
485 490 495
Ser Val Thr Arg Pro Leu Lys Gly His Glu Thr Cys Met Gly Ala Val
500 505 510
Val Ser Glu Tyr Phe Val Leu Thr Ala Ala His Cys Phe Met Val Asp
515 520 525
Asp Gln Lys His Ser Ile Lys Val Ser Val Gly Gly Gln Arg Arg Asp
530 535 540
Leu Glu Ile Glu Glu Val Leu Phe His Pro Lys Tyr Asn Ile Asn Gly
545 550 555 560
Lys Lys Ala Glu Gly Ile Pro Glu Phe Tyr Asp Tyr Asp Val Ala Leu
565 570 575
Val Lys Leu Lys Asn Lys Leu Lys Tyr Gly Gln Thr Leu Arg Pro Ile
580 585 590
Cys Leu Pro Cys Thr Glu Gly Thr Thr Arg Ala Leu Arg Leu Pro Gln
595 600 605
Thr Ala Thr Cys Lys Gln His Lys Glu Gln Leu Leu Pro Val Lys Asp
610 615 620
Val Lys Ala Leu Phe Val Ser Glu Gln Gly Lys Ser Leu Thr Arg Lys
625 630 635 640
Glu Val Tyr Ile Lys Asn Gly Asp Lys Lys Ala Ser Cys Glu Arg Asp
645 650 655
Ala Thr Lys Ala Gln Gly Tyr Glu Lys Val Lys Asp Ala Ser Glu Val
660 665 670

```
Val Thr Pro Arg Phe Leu Cys Thr Gly Gly Val Asp Pro Tyr Ala Asp
        675                 680                 685
Pro Asn Thr Cys Lys Gly Asp Ser Gly Gly Pro Leu Ile Val His Lys
    690                 695                 700
Arg Ser Arg Phe Ile Gln Val Gly Val Ile Ser Trp Gly Val Val Asp
705                 710                 715                 720
Val Cys Arg Asp Gln Arg Arg Gln Gln Leu Val Pro Ser Tyr Ala Arg
                725                 730                 735
Asp Phe His Ile Asn Leu Phe Gln Val Leu Pro Trp Leu Lys Asp Lys
            740                 745                 750
Leu Lys Asp Glu Asp Leu Gly Phe Leu
        755                 760
```

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 737 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

```
Met Thr Ser Met Glu Cys Gly Leu Arg Leu Lys Trp Leu Ile Leu Ala
1               5                   10                  15
Leu Ile Cys Pro Leu Thr Ala Gly Ala Pro Ser Arg Glu Gly Ser Cys
            20                  25                  30
Pro Glu Glu Asn Leu Asp Ile Ala Gly Gly Ser Phe Thr Leu Ser Asn
        35                  40                  45
Gly Tyr Ser Asp Gly Ser Tyr Leu Gln Tyr Ile Cys Pro Asp Asn His
    50                  55                  60
Tyr Pro Ser Ile Ser Ser Arg Arg Cys Gln Phe Gly Val Trp Thr Pro
65                  70                  75                  80
Lys Ala Ser Ser Arg Lys Lys Ala Glu Cys Lys Lys Ile Thr Cys Pro
                85                  90                  95
Asn Pro Arg Val Leu Glu Asn Gly Glu Val Ala Pro Tyr Gln Glu Arg
            100                 105                 110
Tyr Tyr Ile Asn Asp Val Thr Thr Tyr Ser Cys Ser Ser Asp Tyr Lys
        115                 120                 125
Phe Arg Gly Ser Lys Val Arg Val Cys Gln Pro Asn Gly Lys Trp Asn
    130                 135                 140
Gly Ser Thr Pro Ile Cys Gly Arg Asp Ser Asp His Cys Pro Asp Pro
145                 150                 155                 160
Gly Val Pro Pro Gly Ser Ser Arg Thr Gly Ser Ile Phe Asn Ile Asp
                165                 170                 175
Asp Glu Val Thr Tyr His Cys Asp Ser Pro Leu Thr Leu Ile Gly Ser
            180                 185                 190
Lys Val Arg Ser Val Trp Met Tyr Gly Gln Trp Ser Gly Thr Glu Pro
        195                 200                 205
Gln Cys Tyr Ala Asp Phe Thr Tyr Asp Pro Ala Met Glu Ala Ala Glu
    210                 215                 220
Ala Phe Gly Asn Ser Leu Thr Thr Thr Leu Thr Val Gln Gln Gly Phe
225                 230                 235                 240
Glu Asp Asp Gln His Gly Lys Lys Ile Ser Leu Asp Arg Gly Gly Lys
                245                 250                 255
Leu Asp Ile Tyr Ile Ala Val Asp Ala Ser Asp Ser Ile Asp Pro Lys
            260                 265                 270
Asp Phe Asp Lys Ala Lys Lys Ile Ile Lys Thr Leu Ile Glu Lys Ile
        275                 280                 285
Ser Tyr Tyr Glu Val Ser Pro Asn Tyr Glu Ile Leu Met Phe Ala Thr
    290                 295                 300
Asp Val Asp Gln Ile Val Lys Met Arg Asp Phe Lys Thr Asn Glu Lys
305                 310                 315                 320
Ala Arg Lys Ile Leu Lys Ile Phe Glu Asp Leu Asp Asn Phe Asn Tyr
```

```
                    325                                     330                                     335
Asp Lys Lys Gly Asp Arg Thr Gly Thr Asn Ile Ala Lys Leu Tyr Leu
                340                 345                 350
Lys Ile Leu Asp Ser Met Ser Leu Glu Gln Val Gln Asn Lys Glu Asp
            355                 360                 365
Phe Leu Gln Thr Gln His Val Ile Ile Val Phe Thr Asp Gly Gln Ala
    370                 375                 380
Asn Met Gly Gly Asn Pro Lys Pro Lys Val Asp Leu Ile Lys Asn Leu
385                 390                 395                     400
Val Ile Lys Asn Asn Ala Ser Arg Glu Asn Lys Leu Asp Leu Tyr Val
                405                 410                 415
Phe Gly Val Gly Lys Asp Val Lys Lys Glu Asp Met Asn Gly Leu Val
                420                 425                 430
Ser Glu Lys Lys Asp Glu Arg His Phe Phe Lys Leu Pro Asp Leu Asp
            435                 440                 445
Glu Val Gln Asn Thr Phe Asp Leu Met Leu Asp Asp Ser Thr Val Val
    450                 455                 460
Gly Leu Cys Gly Met Gln Gln Asn Tyr Asp Gly Ser Asn Lys Arg Ser
465                 470                 475                     480
Ala Tyr Pro Trp Leu Ala Gln Leu Ser Ile Ala Gln Ser Gln Ile Ser
                485                 490                 495
Asp Cys Met Gly Ser Leu Val Thr Ser Arg Tyr Ile Leu Thr Ala Ala
                500                 505                 510
His Cys Phe Lys Glu Gly Asp Thr Pro Asp Lys Ile Thr Val Tyr Leu
            515                 520                 525
Glu Lys Asn Thr Asp Val Lys Val Glu Lys Val Phe Ile His Pro Asn
    530                 535                 540
Tyr Ser Leu Thr Ala Lys Gln Ser Ile Gly Ile Lys Glu Phe Tyr Asp
545                 550                 555                     560
Phe Asp Val Ala Leu Leu Gln Leu Lys Thr Pro Val Lys Met Ser Val
                565                 570                 575
Asn Leu Arg Pro Ile Cys Leu Pro Cys Thr Lys Glu Thr Asn Arg Ala
                580                 585                 590
Leu Lys Leu Ser Asp Ser Gln Gly Thr Cys Glu Lys His Glu Gln Ile
            595                 600                 605
Leu Leu Ser Asn Glu Leu Val Asp Ala Ala Phe Thr Ser Lys Met Asp
    610                 615                 620
Met Glu Lys Arg Ser Pro Arg Lys Ile Arg Arg Ile Thr Val Lys Leu
625                 630                 635                     640
Gly Lys Tyr Leu Asp Ala Cys Val Glu Asp Ala Lys Lys Ala Lys Glu
                645                 650                 655
Ser Lys Trp Gln Met Arg Arg Arg Gln Leu Gln Lys Ile Ser Cys Gly
                660                 665                 670
Ser Gly Gly Asn Gln Pro Gln Arg Asp Asp Val Ser Cys Lys Gly Glu
            675                 680                 685
Ser Gly Gly Ala Thr His Val Asp Lys Tyr Gly Arg Leu Ile Gln Ile
    690                 695                 700
Gly Val Val Ser Trp Gly Val Lys Asn Leu Cys Ser Lys Lys Arg Asn
705                 710                 715                     720
Leu Met Gln Phe Ser Val Ser Asp Ser Arg Asp Tyr His Ile Asn Pro
                725                 730                 735
Phe
```

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

Asp Ala Thr Met Ser Ile Leu Asp Ile Ser Met Met Thr Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

Asp Ala Thr Met Ser Ile Leu Asp Ile Ser Met Met Thr Gly
1 5 10

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

Asp Gln Asp Ala Thr Met Ser Ile Leu Asp Ile Ser Met Met
1 5 10

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

Ser Ile Leu Asp Ile Ser Met Met Thr Gly Phe Ala Pro Asp Thr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

Lys Ala Phe Ser Asp Arg Asn Thr Leu Ile Ile Tyr Leu Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Glu Val Val Ala Asp Ser Val Trp Val Asp Val Lys Asp Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ser Glu Phe Pro Glu Ser Trp Leu Trp Asn Val Glu Asp Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Leu Ser Ser Asp Phe Trp Gly Glu Lys Pro Asn Leu Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

```
Val Asn Phe Leu Leu Arg Met Asp Arg Ala His Glu Ala Lys
 1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

```
Ala Gln Gly Asp Val Pro Val Thr Val Thr Val His Asp
 1               5                  10
```

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

```
Ser Gly Gln Arg Glu Val Val Ala Asp Ser Val Trp Val Asp Val
 1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

```
Thr Ile Pro Ala Asn Arg Glu Phe Lys Ser Glu Lys Gly Arg
```

1 5 10

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

Ser Ile Thr Val Arg Thr Lys Lys Gln Glu Leu Ser Glu Ala
1 5 10

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

Asp Leu Lys Glu Pro Pro Lys Asn Gly Ile Ser Thr Lys Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

Gly Asp Gly Val Ala Lys Leu Ser Ile Asn Thr His Pro Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO

(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Glu Arg Leu Gly Arg Glu Gly Val Gln Lys Glu Asp Ile
1 5 10

(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

Tyr Tyr Thr Leu Ile Gly Ala Ser Gly Gln Arg Glu Val Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

Asp Gly Ser Pro Ala Tyr Arg Val Pro Val Ala Val Gln Gly Glu
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:68:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Arg Leu Leu Lys Ala Gly Arg Gln Val Arg Glu Pro Gly Gln
1 5 10

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Pro Lys Ser Ser Leu Ser Val Pro Tyr Val Ile Val Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Gln Val Asn Ser Leu Pro Gly Ser Ile Thr Lys Ala Gly Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

Thr Val Leu Thr Pro Ala Thr Asn His Met Gly Asn Val Thr
1 5 10

(2) INFORMATION FOR SEQ ID NO:72:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Glu Val Gln Leu Val Ala His Ser Pro Trp Leu Lys Asp Ser
1 5 10

(2) INFORMATION FOR SEQ ID NO:73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

Thr Ser Asp Leu Asp Pro Ser Lys Ser Val Thr Arg Val Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Thr Ala Glu Leu Val Ser Asp Ser Val Trp Leu Asn Ile Glu
1 5 10

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Met Arg Leu Leu Gly Met Glu Thr Met Ala Trp Gln Glu
1 5 10

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

Arg Glu Ile Leu Asn Ile Asn Gln Lys Arg Asn Asp Tyr
1 5 10

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Trp Arg Val Asn Val Gly Asp Pro Lys Ser Gln Trp Gly Lys
1 5 10

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Leu Lys Thr Ser Ile Gly Asn Lys Pro Pro Glu Lys Leu Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(iii) HYPOTHETICAL: NO
(iv) ANTISENSE: NO
(v) FRAGMENT TYPE: internal
(vi) ORIGINAL SOURCE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

Cys Arg Leu Leu Lys Ala Gly Arg Gln Val Arg Glu Pro Gly Gln Cys
1 5 10 15

I CLAIM:

1. A method for identifying molecules that affect biological activity of a target protein, comprising the steps of:

obtaining information regarding the location of an indel in an amino acid sequence of a target protein;

obtaining a peptide fragment of said target protein, said peptide fragment having a sequence that is located in said amino acid sequence of said target protein within about 30 amino acid residues or less of said indel, or obtaining a peptidomimetic or peptide analog of said peptide fragment; and

screening said peptide fragment, said peptidomimetic, or said peptide analog for its affect on biological or biochemical activity of said target protein.

2. The method of Claim 1, wherein said screening step comprises analyzing for modulation of protein activity, inhibition of protein activity, activation or potentiation of protein activity, competition for binding to a protein, binding to a protein or ligand, substitution for a substrate of said target protein, substitution for a ligand of said target protein, or making an anti-peptide antibody capable of modulating a biological activity of said target protein.

3. The method of Claim 1, wherein said peptide fragment, said peptidomimetic, or said peptide analog directly affects said target protein.

4. The method of Claim 1, wherein said peptide fragment, said peptidomimetic, or said peptide analog indirectly affects said target protein.

5. The method of Claim 1, further comprising the step of synthetically constructing a peptide, peptide analog, or peptidomimetic that affects the biological or biochemical activity of said target protein.

6. A method for making a pharmaceutical composition, comorising the steps of:

obtaining a molecule identified as having biological or biochemical activity in accordance with the method of Claim 1; and

combining said molecule with a pharmaceutically-acceptable carrier.

7. The method of Claim 6, further comprising the step of packaging said molecule in unit-dosage form.

8. The method of Claim 1, wherein said target protein is a protein of the mammalian complement system.

9. The method of Claim 8, wherein said protein of the mammalian complement system is C2, C3, C4, C5 or Factor B.

10. A method of identifying interface peptides for a target protein, comprising:

identifying the location of an indel in an amino acid sequence of a target protein;

selecting an amino acid sequence from said target protein sequence overlapping or located within about 30 amino acid residues or less of an amino- or carboxyl-terminus of an indel;

obtaining a molecule that is a peptide having said selected amino acid sequence, a peptide analog of said selected amino acid sequence, or a peptidomimetic of said selected amino acid sequence; and

evaluating said peptide, peptide analog or peptidomimetic in an assay to measure a change in activity of said target protein, wherein said change in activity is mediated directly or indirectly by said peptide, peptide analog or peptidomimetic.

11. The method of Claim 1 or 10, wherein the peptide fragment has a sequence of about 4 to about 20 amino acid residues, 10 to about 20 amino acid residues, about 4 to about 15 amino acid residues, or about 5 to about 18 amino acid residues in length.

12. The method of Claim 1 or Claim 10, wherein the peptide fragment is located within about 20 amino acid residues, about 15 amino acid residues, about 12 amino acid residues, about 10 amino acid residues, about 8 amino acid residues or less of an indel.

13. The method of Claim 1, wherein said peptide fragment has a sequence that spans said indel.

14. The method of Claim 1, wherein said peptide fragment has a sequence located within said indel.

15. The method of Claim 10, further comprising the step of making antibodies to said peptide, peptide analog or peptidomimetic, wherein said antibodies are capable of modulating an activity of said target protein.

16. The method of Claim 10, wherein the change in activity of said target protein is a decrease in activity, an increase in activity, utilization of a substrate different than the substrate normally utilized by said target protein, or binding to a ligand differently than the ligand binding activity ordinarily demonstrated by said target protein.

17. The method of Claim 10, wherein said target protein is a protein of the mammalian complement system including C2, C3, C4, C5 or Factor B.

18. A peptide for modulating activity of the complement system of a mammal, comprising a sequence of about 4 to about 25 amino acid residues that occurs in an amino acid sequence of a mammalian complement protein, said peptide having an amino acid sequence in which an amino- or carboxyl-terminal residue is located within about 15 amino acid residues of an indel of said mammalian complement protein.

19. The peptide of Claim 18, wherein said peptide modulates activity of the mammalian complement system by directly or indirectly inhibiting an activity of a protein of the mammalian complement system.

20. The peptide of Claim 18, wherein said peptide modulates activity of the mammalian complement system by directly or indirectly enhancing an activity of a protein of the mammalian complement system.

21. The peptide of Claim 18, wherein said indel occurs within the amino acid sequence of a C2, C3, C4, C5 or Factor B protein.

22. The peptide of Claim 18, further comprising another molecule attached to said peptide.

23. An antibody that specifically recognizes a peptide according to Claim 18.

24. A peptide analog or peptidomimetic molecule of a peptide according to Claim 18.

25. A pharmaceutical composition comprising a peptide according to Claim 18.

26. A pharmaceutical composition comprising a peptide analog or peptidomimetic molecule according to Claim 48.

27. A pharmaceutical composition comprising an antibody according to Claim 47.

28. A pharmaceutical composition comprising an antibody according to Claim 49.

29. The peptide of Claim 18, wherein said peptide is located within about 10 amino acid residues of an indel of said mammalian complement protein.

30. A peptide having an amino acid sequence of SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74 or SEQ ID NO:75.

SEQ. ID NO:	PROTIEN	PEPTIDE
1	humC3	VPVAVQGED***********TVQSLTQG*DG
2	musC3	VLVVTQGS************NAKALTQD*DG
3	humC4	IPVKVSATVSSP****GSVPEAQDIQQNTDG
4	musC4	VPVKVSATLVS*****GSDSQVLDIQQSTNG
5	humC5	VPVILNAQTIDVNQETSDLDPSKSVTRVDDG
6	musC5	VPVTLMAQTVDVNQETSDLETKRSITHDTDG

.....<-- INDEL-->.........
N C

<--- divergent region --->

FIG. 1

FIG. 2A

```
                                                          II-1                          III-1
hC3  MGPTSGP--SLLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQ--GDVPVTVTVHDFPGKKLVLSSEKTVLTPATNHMGNVTFTIP
mC3  MGPASGSQLLVLLLLLASSPLALGIPMYSIITPNVLRLESEETIVLEAHDAQ--GDIPVTVTVQDFLKR-QVLTSEKTVLTGASGHLRSVSIKIP
hC4  MRLLWG-----LIWASSFFTLSLQKPRLLLFSPSVVHLGVPLSVGVQLQDVPRGQVVKGSVFLRNPSRNNVPCSPKVDFTLSSERDFALLSLQVP
mC4  MRLLWG-----LAWVFSFCASSLQKPRLLLFSPSVVNLGTPLSVGVQLLDAPPGQEVKGSVFLRNPKGG--SCSPKKDFKLSSGDDFVLLSLEVP
hC5  MGLLG-----ILCFLIFLGKTWGQEQTYVISAPKIFRVGASENIVIQVYGYT--EAFDATISIKSYPDK-KFSYSSGHVHLSSENKFQNSAILTI
mC5  MGLWG-----ILCLLIFLDKTWGQEQTYVISAPKILRVGSSENVVIQVHGYT--EAFDATLSLKSYPDK-KVTFSSGYVNLSPENKFQNAALLTL

         I-1                                         Indel 1
hC3  AWRE-----FKSEKGR--------NK   FVTVQATFGTQVVEKVVLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMVNIENPEGIPVKQDSLSSQNQLGVLPL
mC3  ASKE-----FNSDKEG--------HK   YVTVVANFGETVVEKAVMVSFQSGYLFIQTDKTIYTPGSTVLYRIFTVDNNLLPVGKTVVILIETPDGIPVKRDILSSNNQHGILPL
hC4  LKDAKSCGLHQLLRGPEVQLVAHSP   WLKDSLSRTTNIQGINLLFSSRRGHLFLQTDQPIYNPGQRVRYRVFALDQKMRPSTDTITVMVENSHGLRVR---KKEVYMPSSIFQ
mC4  LEDVRSCGLFDLRRAPHIQLVAQSP   WLRNTAFKATETQGVNLLFSSRRGHIFVQTDQPIYNPGQRVRYRVFALDQKMRPSTDFLTITVENSHGLRVL---KKEIFTSTSIFQ
hC5  QPKQ-----LPGGQNP--------VS   YVYLEVVSKHFSKSKRMPITYDNGFLFIHTDKPVYTPDQSVKVRVYSLNDDLKPAKRETVLTFIDPEGSEV--DMVEEIDHIGIISF
mC5  QPNQ-----VPREESP--------VS   HVYLEVVSKHFSKSKKIPITYNNGILFIHTDKPVYTPDQSVKIRVYSLGDDLKPAKRETVLTFIDPEGSEV--DIVEENDYTGIISF

           Indel 2                                                                                          Indel 3
                                                                                III-12
hC3  S-WDIPELVNMGQWKIRAYYENSPQQVFSTEFE   GSDMVQAERSGIPIVTSPYQIHFTKTPKYFKPGMPFDLMVFVTNPDGSPAYRVPVAVQGED-----------TVQSLTQ
mC3  S-WNIPELVNMGQWKIRAFYEHAPKQIFSAEFE   GSDMVEAERSGIPIVTSPYQIHFTKTPKFFKPAMPFDLMVFVTNPDGSPASKVLVVTQGS------------NAKALTQ
hC4  DDFVIPDISEPGTWKISARFSDGLESNSSTQFE   GGEMEEAELTSWYFVSSPFSLDLSKTKRHLVPGAPFLLQALVREMSGSPASGIPVKVSATVSSP----GSVPEAQDIQQ
mC4  DAFTIPDISEPGTWKISARFSDGLESNRSTHFE   GGEMEEAELTSWRFVSSAFSLDLSRTKRHLVPGAHFLLQALVQEMSGSEASNVPVKVSATLVS-----GSDSQVLDIQQ
hC5  PDFKIPSNPRYGMWTIKAKYKEDFSTTGTAYFE   GGFSEEAEIPGIKYVLSPYKLNLVATPLFLKPGIPYPIKVQVKDSLDQLVGGVPVILNAQTIDVNQETSDLDPSKSVTR
mC5  PDFKIPSNPKYGVWTIKANYKKDFTTTGTAYFE   GGFSEEAEIPGVKYVLSPYTLNLVATPLFVKPGIPFSIKAQVKDSLEQAVGGVPVTLMAQTVDVNQETSDLETKRSITH

      II-3            I-4                                                                A12          Indel 7
hC3  G-DGVAKLSINTHPSQKPLSITVRTKKQELSEAEQATRTMQA       LPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLRMDRAHEAKIRYYTYLIMNKGRL
mC3  D-DGVAKLSINTPNSRQPLTITVRTKKDTLPESRQATKTMEA       HPYSTMHNSNNYLHLSVSRMELKPGDNLNVNFHLRTDPGHEAKIRYYTYLVMNKGKL
hC4  NTDGSGQVSIPIIIPQTISELQLSVSAGSPHP-AIARLTVAA       PP--SGGPGFLSIERPDSRPPR-VGDTLNLNLRAVGS---GATFSHYYYMILSRGQI
mC4  STNGIGQVSISFPIPPTVTELRLLVSAGSLYP-AIARLTVQA       PP--SRGTGFLSIEPLDPRSPS-VGDTFILNLQPVGIP--APTFSHYYYMIISRGQI
hC5  VDDGVASFVLNLPSGVTVLEFNVKTDAPDLPEENQAREGYRA   D5  IAYSSLSQSYLYIDWTDNHKALLVGEHLNIIVTPKSP--YIDKITHYNYLILSKGKI
mC5  DTDGVAVFVLNLPSNVTVLKFEIRTDDPELPEENQASKEYEA       VAYSSLSQSYIYIAWTENYKPMLVGEYLNIMVTPKSP--YIDKITHYNYLILSKGKI
                                                    Indel 8                         Indel 9
```

FIG. 2B

```
     I-5                                   I-6      A14    II-4
hC3  LKAGRQVREPGQDLVVLPLSITTDFIPSFRLVAYYTLIGASGQREVVADSVWVDVKDS-CVGSLV   VKSGQSEDRQPVPGQQMTLKIEGDHGARVVLVAVDKGVF
mC3  LKAGRQVREPGQDLVVLSLPITPEFIPSFRLVAYYTLIGASGQREVVADSVWVDVKDS-CIGTLV   VK-GDPRDNHLAPGQQTTLRIEGNQGARVGLVAVDKGVF
hC4  VFMN--R-EPRKTLTSVSVFVDHHLAPSFYFVAFYYHGDHP-----VANSLRVDVQAGACEGKLE   LS--VDGAKQYRNGESVKLHLETDSLALVALGALDTALY
mC4  MAMG--R-EPRKTVTSVSVLVDHQLAPSFYFVAYFYHQGHP-----VANSLLINIQSRDCEGKLQ   LK--VDGAKEYRNADMMKLRIQTDSKALVALGAVDTALY
hC5  IHFGTREKFSDASYQSINIPVTQNMVPSSRLLVYYIVTGEQ-TAELVSDSVWLNIEEK-CGNQLQ D2 VH-LSPDADAYSPGQTVSLNMATGMDSWVALAAVDSAVY
mC5  VQYGTREKLFSSTYQNINIPVTQNMVPSARLLVYYIVTGEQ-TAELVADAVWINIEEK-CGNQLQ   VH-LSPDEYVYSPGQTVSLDMVTEADSWVALSAVDRAVY
     Indel 10                               Indel 11

hC3  VLNKK--NKLTQSKWDVVE-KADIGCTPGSGKDYAGVFSDAGLTFTSSSGQQTAQRAELQCPQPAARRRR-SVQLTEKRMDK   VGKYP-KELRKCCEDGMRENPMRFSC
mC3  VLNKK--NKLTQSKWDVVE-KADIGCTPGSGKNYAGVFMDAGLAFKTSQGLQTEQRADLECTKPAARRRR-SVQLMERRMDK   AGQYTDKGLRKCCEDGMRDIPMRYSC
hC4  AAGSKSHKPLNMGKFEAMN-SYDLGCGPGGGDSALQVFQAAGLAFS-DGDQWTLSRKRLSCPKEKTTRKKRNVNFQKAINEK   LGQYASPTAKRCCQDGVTRLPMMRSC
mC4  AVGGRSHKPLDMSKFEVIN-SYNVGCGPGGGDDALQVFQDAGLAFS-DGDRLTQTREDLSCPKEKKSRQKRNVNFQKAVSEK   LGQYSSPDAKRCCQDGMTKLPMKRTC
hC5  GVQRG--AKKPLERFQFLE-KSDLGCGAGGGLNNANVFHLAGLTFLTNANADDSQENDEPCKEILRPRR----TLQKKIEEI   AAKYKHSVVKKCCYDGACVN-NDETC
mC5  KVQGN--AKRAMQRFQALDEKSDLGCGAGGGHDNADVFHLAGLTFLTNANADDSHYRDDSCKEILRSKRN-LHLLRQKIEEQ   AAKYKHSVPKKCCYDGARVN-FYETC
     Indel 12

                                                                 II-5          I-8
hC3  QRRTRFISLGEACKKVFLDCCNYITELR-RQHARASHLGLARSNLd--ediiaeenIVSRSEFpeswlwnveDLKEPPKNGISTKLMNIFLKDSI   TTWEILAVSMSDK
mC3  QRRARLITQGENCIKAFIDCCNHITKLR-EQHRRDHVLGLARSELE--EDIIPEEDIISRSHFPQSWLWTIEELKEPEKNGISTKVMNIFLKDSI   TTWEILAVSLSDK
hC4  EQRAARVQQ-PDCREPFLSCCQFAESLRKKSRDKGQAGLQRALEILQEEDLIDEDDIPVRSFFPENWLWRVETVDRF-------QILTLWLPDSL   TTWEIHGLSLSKT
mC4  EQRAARVPQ-QACREPFLSCCKFAEDLRRNQTRSQAHLARNNHNMLQEEDLIDEDDILVRTSFPENWLWRVEPVDSS-------KLLTVWLPDSM   TTWEIHGVSLSKS
hC5  EQRAARISLGPRCIKAFTECCVVASQLRANISHKDMQLGRLHMKTL--LPVSKPE--I-RSYFPESWLWEVHLVPRR-------KQLQFALPDSL   TTWEIQGIGISNT
mC5  EERVARVTIGPLCIRAFNECCTIANKIRKESPHKPVQLGRIHIKTL--LPVMKAD--I-RSYFPESWLWEIHRVPKR-------KQLQVTLPDSL   TTWEIQGIGISDN
                                                      Indel 13                  Indel 14
```

FIG. 2C

```
                                                                                                   II-6
hC3 KGICVADPFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVT-------IPPKSSLSVPYVIVPLKTGLQEVEVKA
mC3 KGICVADPYEIRVMQDFFIDLRLPYSVVRNEQVEIRAVLFNYREQQELKVRVELLHNPAFCSMATAKNRYFQTIK-------IPPKSSVAVPYVIVPLKIGQQEVEVKA
hC4 KGLCVATPVQLRVFREFHLHLRLPMSVRRFEQLELRPVLYNYLDKN-LTVSVHVSPVEGLCLAGG--GGLAQQVL-------VPAGSARPVAFSVVPTAAAAVSLKVVA
mC4 KGLCVAKPTRVRVFRKFHLHLRLPISIRRFEQFELRPVLYNYLNDD-VAVSVHVTPVEGLCLAGG--GMMAQQVT-------VPAGSARPVAFSVVPTAAANVPLKVVA
hC5 -GICVADTVKAKVFKDVFLEMNIPYSVVRGEQIQLKGTVYNYRTSG-MQFCVKMSAVEGICTSESPVIDHQGTKSSKCVRQKVEGSSSHLVTFTVLPLEIGLHNINFSL
mC5 -GICVADTLKAKVFKEVFLEMNIPYSVVRGEQIQLKGTVYNYMTSG-TKFCVKMSAVEGICTSGSSAASLHTSRPSRCVFQRIEGSSSHLVTFTLLPLEIGLHSINFSL
                                                                            Indel 15

                                       I-11
hC3 AVY--HHFISDGVRKSLKVVPEGIRMNKTVAVrTLDPERLGreqvqkedippADLSDQVPDTESETRILLQGTPVAQMT--EDAVDAERLKHLIVTPSGCGEQNMIG
mC3 AVF--NHFISDGVKKTLKVVPEGMRINKTVAIHTLDPEKLGQGGVQKVDVPAADLSDQVPDTDSETRIILQGSPVVQMA--EDAVDGERLKHLIVTPAGCGEQNMIG
hC4 RGSFE-FPVGDAVSKVLQIEKEGAIH-REELVYELNPLDHRG---RTLEIPGNSDPNMIPDGDFNSYVRVTASDPLDTLGSEGALSPGGVASLLRLPRGCGEQTMIY
mC4 RG--V-FDLGDAVSKILQIEKEGAIH-REELVYNLDPLNNLG---RTLEIPGSSDPNIVPDGDFSSLVRVTASEPLETMGSEGALSPGGVASLLRLPQGCAEQTMIY
hC5 ETW----FGKEILVKTLRVVPEGVKR-ESYSGVTLDPRGIYGTISRRKEFPYRIPLDLVPKTEIKRILSVKGLLVGEIL--SAVLSQEGINILTHLPKGSAEAELMS
mC5 ETS----FGKDILVKTLRVVPEGVKR-ESYAGVILDPKGIRGIVNRRKEFPYRIPLDLVPKTKVERILSVKGLLVGEFL--STVLSKEGINILTHLPKGSAEAELMS
                          Indel 16                                  Indel 17

hC3 MTPTVIAVHYLDETE   QWEKFG---LEKRQGALELIKKGYTQQLAFRQPSSAFAAFVKRAPSTWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVFQ
mC3 MTPTVIAVHYLDQTE   QWEKFG---IEKRQEALELIKKGYTQQLAFKQPSSAYAAFNNRPPSTWLTAYVVKVFSLAANLIAIDSHVLCGAVKWLILEKQKPDGVFQ
hC4 LAPTLAASRYLDKTE   QWSTLP---PETKDHAVDLIQKGYMRIQQFRKADGSYAAWLSRDSSTWLTAFVLKVLSLAQEQVGGSPEKLQETSNWLLSQ-QQADGSFQ
mC4 LAPTLTASNYLDRTE   QWSKLS---PETKDHAVDLIQKGYMRIQQFRKNDGSFGAWLHRDSSTWLTAFVLKILSLAQEQVGNSPEKLQETASWLLAQ-QLGDGSFH
hC5 VVPVFYVFHYLETGN   HWNIFHSDPLIEKQKLKKKLKEGMLSIMSYRNADYSYSVWKGGSASTWLTAFALRVLGQVNKYVEQNQNSICNSLLWLVENYQLDNGSFK
mC5 IAPVFYVFHYLEAGN   HWNIFYPDTLSKRQSLEKKIKQGVVSVMSYRNADYSYSMWKGASASTWLTAFALRVLGQVAKYVKQDENSICNSLLWLVEKCQLENGSFK
                      Indel 18
```

FIG. 2D

```
                                                   II-7
hC3 EDAPVIHQEMIGGLRNNN-EKDMALTAFVLIS LQEAKDICEEQ-----VNSLPGSITKAGDFLEANYMN-LQRSYTVAIAGYALAQMGRLKGPLLNKflttakdkn
mC3 EDGPVIHQEMIGGFRNAK-EADVSLTAFVLIA LQEARDICEGQ-----VNSLPGSINKAGEYIEASYMN-LQRPYTVAIAGYALALMNKLEEPYLGKFLNTAKDRN
hC4 DPCPVLDRSMQGGLVGND--ETVALTAFVTIA LHHGLAVFQDEGAEPLKQRVEASISKANSFLGEKASAGLLGAHAAAITAYALSLTKAPVDLLGVAHNNLMAMAQ
mC4 DPCPVIHRAMQGGLVGSD--ETVALTAFVVIA LHHGLDVFQDDDAKQLKNRVEASITKANSFLGQKASAGLLGAHAAAITAYALTLTKASEDLRNVAHNSLMAMAE
hC5 ENSQYQPIKLQGTLPVEARENSLYLTAFTVIG IRKAFDICP-------LVKIDTALIKADNFLLENTLP-AQSTFTLAISAYALSLGDKTHPQFRSIVSALKREAL
mC5 ENSQYLPIKLQGTLPAEAQEKTLYLTAFSVIG IRKAVDICP-------TMKIHTALDKADSFLLENTLP-SKSTFTLAIVAYALSLGDRTHPRFRLIVSALRKEAF
                    Indel 19                 Indel 20

hC3 r---------wed------------pgkqlyn----------veats yallallqlkdfd-fvppvvrwlneqryyggygstqATFMVFQALAQYQKDAPDHQELN
mC3 R---------WEE------------PDQQLYN----------VEATS YALLALLLLKDFD-SVPPVVRWLNEQRYYGGGYGSTQATFMVFQALAQYQTDVPDHKDLN
hC4 ETGDN---LYWGSVIGsQSNAVSPTPAPRNPSDPMPQAPALWIETTA YALLHLLLHEGKAEMADQASAWLTRQGSFQGGFRSTQDTVIALDALSAYWIASHTTEERG
mC4 ETGEH---LYWGLVLGSQDKVVLRPTAPRSPTEPVPQAPALWIETTA YALLHLLLREGKGKMADKAASWLTHQGSFHGAFRSTQDTVVTLDALSAYWIASHTTEEKA
hC5 VKGNPPIYRFWKDNL---------QHKDSSVPNT----GTARMVETTA YALLTSLNLKDIN-YVNPVIKWLSEEQRYGGGFYSTQDTINAIEGLTEYSLLVKQLR-LS
mC5 VKGDPPIYRYWRDTL---------KRPDSSVPSS----GTAGMVETTA YALLASLKLKDMN-YANPIIKWLSEEQRYGGGFYSTQDTINAIEGLTEYSLLLKQIH-LD
                    Indel 21

hC3 LDVSLQLPSrSSKITHRIHWESA-SLLrSEETKENEGFTVTAE--GKGQGTLSVVTMYHAK
mC3 MDVSFHLPSRSSATTFRLLWENG-NLLRSEETKQNEAFSLTAK--GKGRGTLSVVAVYHAK
hC4 LnVTLSSTGrNGFKSHALQLNNRQIRGLEEELQFSLGSKINVKVGGNSKGTLKVLRTYNVL
mC4 LkVTLSSMGRNGLKTHGLHLNNHQVKGLEEELKFSLGSTISVKVEGNSKGTLKILRTYNVL
hC5 MDIDVSYKHKGALHNYKMTDKN--FLGRPVEVLLNDDLIVSTGF-GSGLATVHVTTVVHKT
mC5 MDINVAYKHEGDFHKYKVTEKH--FLGRPVEVSLNDDLVVSTGY-SSGLATVYVKTVVHKI
             Indel 22        Indel 23
                                                                                           II-8A/III-8B/III-8C
hC3 AKDQLTCNKFDLKVTIKPAPETEKRP----------------------------------------QDAKNTMILEICTRYRGDQD-----ATMSILDISM
mC3 LKSKVTCKKFDLRVSIRPAPETAKKP----------------------------------------EEAKNTMFLEICTKYLGDVD-----ATMSILDISM
hC4 DMKNTTCQDLQIEVTVKGHVEYTMEANEDYEDYEYDELPAKDDPDAPLQPVTPLQLFEGRRNRRRREAPKVVEEQESRVHYTVCIWRNGKVG----LSGMAIADVTL
mC4 DMKNTTCQDLQIEVKVTGAVEYAWDANEDYEDYY--DMPAADDPSVPLQPVTPLQLFEGRRSRRRREAPKVAEEQESRVQYTVCIWRNGKLG----LSGMAIADITL
hC5 STSEEVCS-FYLKIDTQDIEASHYRG----------------------------------------YGNSDYKRIVACASYKPSREESSSGSSHAVMDISL
mC5 SVSEEFCS-FYLKIDTQDIEASSHFR----------------------------------------LSDSGFKRIIACASYKPSKEESTSGSSHAVMDISL
                         Indel 24                                                           Indel 25
```

FIG. 2E

III-11

hC3 MTGFAPDTDDLKQLANGV DRYISKYELDKAFSDRNTliivldkvshsEDDCLAFKVHQYFNVELIQPGAVKVYAYYNLEESCTRFYHPEKEDGKLNKLCRDELCRCAEENC
mC3 MTGFAPDTKDLELLASGV DRYISKYEMNKAFSNKNTLIIYLEKISHTEEDCLTFKVHQYFNVGLIQPGSVKVYSYYNLEESCTRFYHPEKDDGMLSKLCHSEMCRCAEENC
hC4 LSGFHALRADLEKLTSLS DRYVSHFETEGPH-----VLLYFDSVPTSR-ECVGFEAVQEVPVGLVQPASATLYDYYNPERRCSVFYGAPSKSRLLATLCSAEVCQCAEGKC
mC4 LSGFHALRADLEKLTSLS DRYVSHFETDGPH-----VLLYFDSVPTTR-ECVGFGASQEVVVGLVQPSSAVLYDYYSPDHKCSVFYAAPTKSQLLATLCSGDVCQCAEGKC
hC5 PTGISANEEDLKALVEGV DQLFTDYQIKDGH-----VILQLNSIPSSDFLCVRFRIFELFEVGFLSPATFTVYEYHRPDKQCTMFYSTS--NIKIQKVCEGAACKCVEADC
mC5 PTGIGANEEDLRALVEGV DQLLTDYQIKDGH-----VILQLNSIPSRDFLCVRFRIFELFQVGFLNPATFTVYEYHRPDKQCTMIYSIS--DTRLQKVCEGAACTCVEADC

Indel 26 Indel 27 A15

hC3 FIQKS--DDKVTLEE--RLDKACE-PGVDY VYKTRLVKVQLSNDFDEYIMAIEQTIKSGSDEVQV-GQQRTFISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSYI
mC3 FMQQS--QEKINLNV--RLDKACE-PGVDY VYKTELTNIKLLDDFDEYTMTIQQVIKSGSDEVQA-GQQRKFISHIKCRNALKLQKGKKYLMWGLSSDLWGEKPNTSYI
hC4 PRQRRALERGLQDEDGYRMKFACYYPRVEY GFQVKVLREDSRAAFRLFETKITQVLHFTKDVKAANQMRNFLVRASCRLRLEPG--KEYLIMGLDGATYDLEGHPQYL
mC4 PRLLRSLERRVEDKDGYRMRFACYYPRVEY GFTVKVLREDGRAAFRLFESKITQVLHFRKDTMASIGQTRNFLSRASCRLRLEPN--KEYLIMGMDGETSDNKGDPQYL
hC5 GQMQEELDLTISAET--RKQTACK-PEIAY AYKVSITSITVENVFVKYKATLLDIYKTGEAVAEK-DSEITFIKKVTCTNAELVK-GRQYLIMGKEALQIKYNFSFRYI
mC5 AQLQAEVDLAISADS--RKEKACK-PETAY AYKVRITSATEENVFVKYTATLLVTYKTGEAADE--NSEVTFIKKMSCTNANLVK-GKQYLIMGKEVLQIKHNFSFKYI

Indel 28 Indel 29

hC3 IGKD--TWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN
mC3 IGKD--TWVEHWPEAEECQDQKYQKQCEELGAFTESMVVYGCPN
hC4 LDSN--SWIEEMPSERLCRSTRQRAACAQLNDFLQEYGTQGCQV
mC4 LDSN--TWIEEMPSEQMCKSTRHRAACFQLKDFLMEFSSRGCQV
hC5 YPLDSLTWIEYWPRDTTCS--SCQAFLANLDEFAEDIFLNGC
mC5 YPLDSSTWIEYWPTDTTCP--SCQAFVENLNNFAEDLFLNSCE

Indel 30 Indel 31

3A

```
HSV-1   MFSGGGGPLSPGGKSAARAASGFFAPAGPRGAGR-GPPPCLRQNFYNPYLAPVGTQQKPTGPTQRHTYYSECDEFRFIAPRVLDEDAPPEKRAGVHD
HSV-2   MFCAAGGPASPGGKSAARAASGFFAPHNPRGATQTAPPPCRRQNFYNPHLAQTGTQPKAPGPAQRHTYYSECDEFRFIAPRSLDEDAPAEQRTGVHD
HSV-6                                          MDSVSFFNPYLEANRLKKK-----SRSSYIR-------ILPRGIMHDGAAGLIKDVCD
Baculo                                         MSSVNLMEWSALKTQLQAG---RDAGKARVS-------IGP----ADTARITRMTYAD
CCV                                  MDRNAVLYGVLEHRLPKWVELSDDTDLEPF---FFSSVRYIT-------AGS----EDAIMIQALNLNT
ENTPOX                                  MKLKKLYIFYFDIYEYFLCDLQLSET----NEILKYIKNN-----IDKY---TNSFNSSYIILKD
HepB                                       MPLSYQHFRKLLLLDDETEAG--------------------P----LEE-ELPRLADAD

HSV-1   GHLKRAPKVYCGGDERDVLRVGS GGFWPRRSRLWGGVDHAPAG-FNPTVTVFHVYDILENVEHAYGMRAAQFHARFM-DAITPTGTVITLLGLTPEG
HSV-2   GRLRRAPKVYCGGDERDVLRVGP EGFWPRRLRLWGGADHAPEG-FDPTVTVFHVYDILEHVEHAYSMRAAQLHERFM-DAITPAGTVITLLGLTPEG
HSV-6   ----SEPRMFYR-D-RQYLLS-K EMTWP-------SLDIARSKDYDHMRMKFHIYDAVETLMFTDSIENLPFQYR---HFVIPSGTVIRMFGRTEDG
Baculo  ----NHLIVFMN----ARLAK---------------ENHRLYQFYAEVRCDLYSYKSCYGTHASATCHRNCTSYK---TFVMPGLRDV----HTDKL
CCV     ----DEIVVFLV----TNLNF--------------MALIPTVYIENPGIRQLIASTPISYRSPITVFNGDLKKWMDCDLFVFGTMAAQK--AFIKAG
ENTPOX  ----FNIITNEV-ELQSYYNF--------------TEDSKIKL-NNTDLILFMTPYKIERIYSKYNRNFNQYRWFYILNNIEPAGSYK-----INMS
HepB    ----LNRRVAED------LNL---------------GNLNVSIPWTHKVGNFTGLY-------SSTVPIFNPEWQ---TPSFPKIHLH------EDI

HSV-1   HRVAVHVYGTRQYFYMNKEEVDRHLQCRAPRDLCERMAAALRESPG ASFRGISADHFEAEVVERTDVYYYETRPALFYRVYVR-SGRVLSYLCDNFC
HSV-2   HRVAVHVYGTRQYFYMNKAEVDRHLQCRAPRDLCERLAAALRESPG ASFRGISADHFEAEVVERADVYYYETRPTLYYRVFVR-SGRALAYLCDNFC
HSV-6   EKICVNVFGQEQYFYC--ECVD----GRSLKATINNLMLTGEVKMS CSFVIEPADKLSLYGYNANTVVNLFKVSFGNFYVSQR-IGKILQ------N
Baculo  HVVKFKRS--DEKRDK--NCLD----GYLADVNRVHMQTSLLEGQY VRFKNAHACRDYRLSHTAKDVHEFESM-LERVQVSAL-SHEILP-------
CCV     NSVLGSLGG-NVYTYG--DHVS----NFDGNTPVLQNNLMCSHVYY TRYKTDVYAPWEFYYDQKRDQGYLMSLPAIIPRCKREGAFDIET-------
ENTPOX  NLQNINIYDKNKTAYY---CKN----PKLLFLTPIEIDKFIPVSR- VSIDIECQHFGEFPTPNKFPISHICIDWFMESNINP--VKKIITLIN---Y
HepB    ANRCQQFVG--PLTVN---EKR-----RLKLIMPARFYPNSTKYLP LDKGIKTYYPDHVVNHYFQTRHYLHTLWKAGILYKR----ETTR-------

HSV-1   PAIKKYEGGVDAT-TRFILDNP-GFVTFGWYRLKPG-RNNTLAQPRAPMAFGTSSDVEFNCTADNLAIE GGMSDLPAYKLMCFDIECKAGGEDELAF
HSV-2   PAIRKYEGGVDAT-TRFILDNP-GFVTFGWYRLKPG-RGNAPAQPRPPTAFGTSSDVEFNCTADNLAVE GAMCDLPAYKLMCFDIECKAGGEDELAF
HSV-6   EGFVVYEIDVDVL-TRFFVDN--GFLSFGWYNVKKY-IPQDMGK-------GSNLEVEINCHVSDLVSL EDVN-WPLYGCWSFDIECLGQNGN---F
Baculo  -VVACYDIETHSDGQRFSAPDA-DFIISIAVVVRRD-AADTRIC------LFYSPDDPVDLSSS--SSS PPAA--PDTAAVHFRAERDMIAAFFQLL
CCV     -IVHENAMDQDLNCQKFFKSEF-RSMEESQVLIQRFREAGVTGLPPSP-FVGITQKLHEIVSISLVVCN YHKTG-PKKKEY-YVYYNTKKMEN----
ENTPOX  EIIKNYKGEQK-D--RFIYTEIDELLTKDKVYITIY-CTEKYMLH----FILYTLRKDFDYILTYNGHS FDFT--YIQGRRKFYNLNELCLVN---A
HepB    --SASFCGS---P---YSWEQ---ELHHGRLVIKTS---QRHG------------DEPFCSQPSGILSR SSVG--PCIRSQ-FKQSRLGLQPH--QG
```

FIG.3B

```
HSV-1   PVAGHP-EDLVIQISCLLYDLSTTALEHVLLFSLGSCDLPESHLNELAARGLPTPVVLEFDSEFEMLLAFMTLVKQYGPEFVTGYN-IINFD
HSV-2   PVAERP-EDLVIQISCLLYDLSTTALEHILLFSLGSCDLPESHLSDLASRGLPAPVVLEFDSEFEMLLAFMTFVKQYGPEFVTGYN-IINFD
HSV-6   PDAENL-GDIVIQISVISFDTEGDRDER-HLFTLGTCEKIDGVH-----------IYEFASEFELLLGFFIFLRIESPEFITGYN-INNFD
Baculo  PLLN---ADVVLDFNGDKFDLPFL/TG-----RANKLCGPAEAARATKIAR------YDLSPVNVVTQQSYDKFSNKLHSHYL/TYYIHIDLYQ
CCV     PMEMIPVEHLHLDASRIKFEACKN-------EFYMLLAFINRLR-------------KSVNVLYVYNAQFDIQVIQQR---LRYYA--FKQR
ENTPOX  HKSN-E-LKIYSYNKDTTYEIDSNNG----IIFLDLYNYIKKIYN-----------YNSYKLGBIAKERFNILSKIID--NGDEYI-IMPLD
HepB    PLATS--QPGRSGSIWARVHSPTR-------RCFGVEPSGSGHIG-----------HRASDASSCLHQS-AVRKAAYS-----HLS--TSKR

HSV-1   WPFLLAKI/TDIYKVPLDGYGRMN-GRG--VFRVWDIGQSHFQKRSKIKVNGMVNIDMYGI-ITDKIKLSSYKLNAVAEAVLKDKKKDLSYRDIPAYYATGPAQRGVIGEY
HSV-2   WPFVL/TKL/TEIYKVPLDGYGRMN-GRG--VFRVWDIGQSHFQKRSKIKVNGMVNIDMYGI-ITDKVKLSSYKLNAVAEAVLKDKKKDLSYRDIPAYYASGPAQRGVIGEY
HSV-6   LKYLCIRMDKIYHYDIGCFSKLKNGKIGISVPHEQYRKGFLQAQTKVFTSGVLYLDMYPV-YSSKITAQNYKLDTIAKICLQQEKEQLSYKEIPKKFISGPSGRAVVGKY
Baculo  FLSTDSEHNDLENFQLNTVAEHYLKKSKVDLPIHDMLQMYGEKRLSRIVEYNVQDCVLPVELFLKLEIADYMYTQCMLLYLCT---DDLLRNISHKITVAYFHLAL/TNTV
CCV     APRCCKGHDDIPHEWGKALMEKWEAFL--SVKPQLFKAQILMGQDILKANYLKLLEGIGS-VLAQAKSTMAKMCTIKERIDSYRKMKDTVQNFKS-HGFGCDIID-MMYV
ENTPOX  TADNKNKVSIFYDV-IRTANYCFINNI--PYKIKDKTKIINDKE-KLYDPISIENSLYQQ--FKIYKNNTPISDETTKVMLSKD--DVDIGNKNAYVNFTKDKSDDIAYY
HepB    QSS--SGHAVEFHS---------------FPPSSARSQS-QGP--VFSCWWLQFRNTQP-------CSNYCLSHLVNLLEDWG------PCTEH----GEHHIR----I

HSV-1   CIQDSLLVGQ LFFKFLPHLELSAVARLAGINITRTIYDGQQIRVFTCLLRLADQKGFILPDTQGRFRGAGGEAPKRPAAAREDEERP-----EEEGEDEDEREEGGKRE
HSV-2   CVQDSLLVGQ LFFKFLPHLELSAVARLAGINITRTIYDGQQIRVFTCLLRLAGQKGFILPDTQGRFRGLDKEAPKRPAVPRGEGERPGDGNGDEDKDDDEDGDEDGDERE
HSV-6      DSVLVVR LFKQINYHFEVAEVARLAHVTARCVVFEGQQKKIFPCII/TEAKRRNMILPSMVSSHN--------------------------------------RQ
Baculo  ARRPDPTPDP YFFN-KYDLSVTSGASAPSTSRPANAIDLSQLKRTP---VDAAR---IPPSAVKLCS--------------------------------------TRQ
CCV     CKRKEFEAKD GSLNTVAQLIIKKFKPHKATPKIHKMDDITYDKLDG--YYRAGG-TKIAECLIYNLI-------------------------------------DSL
ENTPOX  CTHDTVLCNC IFKYDMIHDKVIAFSNEYLLPQYMSFKYKSTTNISG--LLLKTL--FCNRSMIVSGN-------------------------------------LEFE
HepB    PRTPARVTGG VFLV----D-----KNPHNTAESRLVVDFSQFSRGS---TRVSWPKFAVPNLQSL/TN---------------------------------------

HSV-1   PEGARETAGRHVGYQGAKVL DPTSGFHVNPVVVFDFASLYPSIIQAHNLC-FSTLSLRADAVAHLEAGKDYLEIEVGGRRLFFVKAHVRESLLSILLRDWLAMRKQIRSRI
HSV-2   -EVARETGGRHVGYQGARVL DPTSGFHVDPVVVFDFASLYPSIIQAHNLC-FSTLSLRPEAVAHLEADRDYLEIEVGGRRLFFVKAHVRESLLSILLRDWLAMRKQIRSRI
HSV-6   --------G--IGYKGATVL EPKTGYYAVPTVVFDFQSLYPSIMMAHNLC-YSTLVLDERQIAGLSES-DIL/TVKLGDETHRFVKPCIRESVLGSLLKDWLAKRREVKAEM
Baculo  --------S--CTYKGGKVL SPKPGFNR-WVATLDFNALYPTIMMWEGVC-MSSNVFIAS------DGNVYLDKNVNAVNPKLLKTLSEMRVRYKGLRDQCEYN---SFYY
CCV     ----------LVIRIAKNL KPMEEYTYRQLACYNIDTAAHTRG-VMNFCGFIQSTKVVEVS----RNKARLDAGIVMATDYIRNSLFTPETIPR--RGGFVMAPLTGLFF
ENTPOX  ------------KFEGGYVL EPKQKYIDSITAVFDFNSEYPSNIIEANLS-PEKVEKVIK-----LQD-DEYAVDIVEN--YLKEKYPYPDYCYMLIKKDKTYKFIVMDRR
HepB    --------------LLSSNL S---WLSLDVSAAFYHIPLHPAAM---------PHLLIGS------SGLSRYVARLSSN---SRINNNQHGTLQN-LHDSCSRQ----LYV
                          *
```

FIG. 3C

```
HSV-1   P-QSSPEEAVLLDKQQAAIKVVCNSVYGF TGVQHGLLPCLHVAATVTTIGREMLLATREYVHARWAAFEQLLADFPEAADMRAPGPYSMRIIYGDTDSIFVLCRG
HSV-2   P-QSPPEEAVLLDKQQAAIKVVCNSVYGF TGVQHGLLPCLHVAATVTTIGREMLLATRAYVHARWAEFDQLLADFPEAAGMRAPGPYSMRIIYGDTDSIFVLCRG
HSV-6   QNCSDPMMKLLLDKKQLALKTTCNSVYGV TGAAHGLLPCVAIAASVTCLGREMLCSTVDYVNSKMQS-EQFFCEEFGLTSSDFTGDLEVEVIYGDTDSIFMSVRN
Baculo  K---------LYDKIQNALKRIANSIYGY YGIF--FKP---LANYITKMGRGKLKEVVGKVEAMSDD--PRILREFGLSKINFS------VIYGDTDSCFIRVLF
CCV     A---RP-----TQCFELCLD--FTSMYPS MMCDLNISP-----ETIVDS---DKTNRVGDYMGYDWSKIDQGFEKFTLVLRVDRTDPENPKLVRHTSDTSLSLKRY
ENTPOX  K-----P--GIITQMIDKGMK-SKNEYKNL KNIN-KNNP--VLYNYYTSALYSKKITINSLYGLLGSE--RFDFNSPYCAEYCTA---------LGQKCIKYIKNLV
HepB    S-----------LMLL----YK-----TYGW KLHLY-SHP---------IILGFRKIPMGVGLSPFLLAQ---FTSAICSVVRRAFP---------H-CLAFSYMDDVV
                                          *

HSV-1   LTAAGLTAMGDK--MASHISRALFLP--------PIKLECEKTFT KLLLIAKKKYIGVIYGGKMLI-KGVDLVRKNNCAFINRTSRALVDLLFYDDTVSGAAAALAE
HSV-2   LTGEALVAMGDK--MASHISRALFLP--------PIKLECEKTFT KLLLIAKKKYIGVICGGKMLI-KGVDLVRKNNCAFINRTSRALVDLLFYDDTVSGAAAALAE
HSV-6   MVNQSLRRIAPM--IAKHITDRLFKS--------PIKLEFEKILC PLILICKKRYIGRQDDSLLIF-KGVDLVRKTSCDFVKGVVKDIVDLLFFDEEVQTAAVEFSH
Baculo  DEAEWRRTAARP--RSAPSCRTTCAKRSTTLWCGYKMSLENIML SLILLKKKKYCYLNNEQRTKY-KGWLIKR-DMPLFMRKAFRATVDS-FSAATRRVRARPARR
CCV     LRLRTEHKRALK--QSSGSVAEYHNR-------LQNEMKICTN-T HYGVSEHTCSLMITTTQGQHKI-KLVNEFIKTLNRTGHSLFPNYGDT---DSTMLYHPSDESE
ENTPOX  DKSRYIDNNLYLNEQNNPFSNEPVIT---------RYSGNLDVNFT FYIIYGDTDSLFINIKFDNKFDNKEDLVNKSHECFQF-LSNIINDE---KNIILSKNFNFEY
HepB    LGAKSVQHLESL---YTAVTN-FLLS---------LGIHLNPNKT KRWGYSLN-FMGYVIGSWGTL-PQDHIVQKIKHCFRKLPVNRPIDW------KVCQRLVGL-L
                                                                                                         *

HSV-1   RPAEEWLARPLP--EGLQAFGAVLVDAHRRITDP-ERDIQDFVLTAELSRHPRAYTNK RLAHLTVYYKLMARRAQVPSIKDRIP-YVIVAQTREVEETVARLAALRE
HSV-2   RPAEEWLARPLP--EGLQAFGAVLVDAHRRITDP-ERDIQDFVLTAELSRHPRAYTNK RLAHLTVYYKLMARRAQVPSIKDRIP-YVIVAQTREVEETVARLAALRE
HSV-6   MTQTQLREQGVP--VGIHKILRRLCEAREELFQN-RADVRHLMLSSVLSKEMAAYKQP NLAHLSVIRRLAQRKEEIPNVGDRIM-YVLIAPS---------------
Baculo  EMLRYYREFGAPR-ENLVDYCFSMSYNETSTTAK-RRKEED-------PARKPVITIAK ---HCRELLANPG-VDFLPGNGDRIQ-YVLVDVK---EKITQKAFPLKL
CCV     TQLEDMVTLEDFMRAELREYMLKKLSAELVNRVKEKTKRTDTFVQSFLSDVETVLFDD ---MVEKLRLFSQ-GEVIEPFKDGGT-WVVVDPL---------------
ENTPOX  EKMYTWMLLLAPK-KYIGEVVSSMNPLQLISDSKGTALIRR----DCTEIHKTILKNT ----IDILKEYLTNNCTIQDVNNKINNYLMFTFK---------------
HepB    GFAAPFTQCGYP----ALMPLYACIQAKQAFTFSP-----TYK--------AFLSKQYMNL ----YPVARQRPG-LCQV--FADATPTGWGLAIG---------------

HSV-1   LDAAAPGDEPAPPAALPSPAKRPRETPSHADPPGGASKPRKLLVSELAEDPAYAIAHGVALNTDYYFSHLL GAACVTFKALFGNNAKITES-LLKRFIPEVWHPPD
HSV-2   LDAAAPGDEPAPPAALPSPAKRPRETPSHADPPGGASKPRKLLVSELAEDPGYAIARGVPLNTDYYFSHLL GAACVTFKALFGNNAKITES-LLKRFIPETWHPPD
HSV-6   ------IGN------------KQ-------THN---------------YELAEDPNYVIEHKIPIHAEKYFDQII KAVTNAISPIFPKTDIKKEK-LLLYLLPMKVYLDE
Baculo  FD---PDS----PTLQISWLKHMNILCTFMNELIQVFGNRPEFEHYFGAIVDEYTSAQMYDVRYPVLVPTRR AKAGKSAKKNDSDSDSDSDDDDDPATTPVNYH--S
CCV     ------TGI----------WMDC-------STP-------------FS-SELICKLEYENASSIGCHVAKKMVSIG STYLF-FKKISLYHVRVWRM--CADTDGSPSH--L
ENTPOX  --------N-----------IIEN--------IQN-------------LDINEFKKSVKYTQIYKDPN---FYIELCV KKYN---SKN-P-NDKIVKG---QRFDFIYAHEID
HepB    ------------------------------HQ-------------------R--MRETFVAPLPIHTAELLAACF ARSRS-GAKLIGTDNSVVLS---QKYTS---FP
```

FIG. 3D

```
HSV-1    DVAARLRAAGFGAVGAGATAEETRRMLHRAFDTLA
HSV-2    DVAARLRAAGFGPAGAGATAEETRRMLHRAFDTLA
HSV-6    TFSAIAEVM
Baculo   LFSMHLKKPKRQAVG---EFEPCPQCVARA
CCV      YFPVSLSRTRAKQRG-----------DH
ENTPOX
HepB     IWDIETKKWNTKYTS
```

FIG. 4A

```
HuC3    MGPTSGP--SLLLLLLTHLPLALGSPMYSIITPNILRLESEETMVLEAHDAQGDVPVTVTVHDFPGKKLVLSSEKTVLTPATNHMGNVTFTIPANREFKSEK
MuC3    MGPASGSQLLVLLLLLASSPLALGIPMYSIITPNVLRLESEETIVLEAHDAQGDIPVTVTVQDF-LKRQVLTSEKTVLTGASGHLRSVSIKIPASKEFNSDK
RATC3   MGPTSGSQLLVLLLLLASSLLALGSPMYSIITPNVLRLESEETFILEAHDAQGDVPVTVTVQDF-LKKQVLTSEKTVLTGATGHLNRVFIKIPASKEFNADK
GUIPIG  MGPAAGP--SLLLLLLASVSLALGDPMYSIITPNILRLENEETVVLEAHEVQGDIPVTVTVHDFPAKKNVLSSEKTVLTSATGYLGTVTIKIPASKEFKSDK
        ***..*   .*****.   **** *********.**** *** .****. ***.******.**  *. **.******* *.. .   *   *** .** ..*

HuC3    GRNKFVTVQATFGTQVVE KVVLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVNHKLLPVGRTVMVNIENPEGIPVKQDSLSSQNQLGVLPLSWDIPELVNM
MuC3    EGHKYVTVVANFGETVVE KAVMVSFQSGYLFIQTDKTIYTPGSTVLYRIFTVDNNLLPVGKTVVILIETPDGIPVKRDILSSNNQHGILPLSWNIPELVNM
RATC3   -GHKYVTVVANFGATVVE KAVLVSFQSGYLFIQTDKTIYTPGSTVFYRIFTVDNNLLPVGKTVVIVIETPDGVPIKRDILSSHNQYGILPLSWNIPELVNM
GUIPIG  -GRKLVVVQAAFGGTQLE KVVLVSLQSGYLFIQTDKTIYTPGSTVLYRIFTVDSDLLPVGRTIIVTIETPDGIPIKRDTLSSNNQHGILPLSWNIPELVNM
          * * * * * **  .* * *.** ******************** ******   *****.*... **.*.*.*.*.* *** ** *.***** *******

HuC3    GQWKIRAYYENSPQQVFSTEFEVKEYVLPSFEVIVEP TEKFYYIYNEKGLEVTITARFLYGKKVEGTAFVIFGIQDGEQRISLPESLKRIPIEDGSGEVVLS
MuC3    GQWKIRAFYEHAPKQIFSAEFEVKEYVLPSFEVRVEP TETFYYIDDPNGLEVSIIAKFLYGKNVDGTAFVIFGVQDGDKKISLAHSLTRVVIEDGVGDAVLT
RATC3   GQWKIRAFYEHAPKQTFSAEFEVKEYVLPSFEVLVEP TEKFYYIHGPKGLEVSITARFLYGKNVDGTAFVIFGVQDEDKKISLALSLTRVLIEDGSGEAVLS
GUIPIG  GQWKIQAFYENSPKQVFSAEFEVKEYVLPSFEVLVEP TEKFYYIDDPKGLEVNIIARFLYGKNVDGTAFVIFGVQDGDQRISLAQSLTRVVIEDGSGEVVLS
        *****.*.**..*.* **.*************** *** ** ****     **** * *.***** *.********.** ...*** ** *. **** *. **.

HuC3    RKVLLDGVQNLRAEDLVGKSLYVSATVILHSGSDMVQAERSGIPIVTSPYQIHFT KTPKYFKPGMPFDLMVFVTNPDGSPAYRVPVAVQGEDTVQSLTQGDG
MuC3    RKVLMEGVRPSNADALVGKSLYVSVTVILHSGSDMVEAERSGIPIVTSPYQIHFT KTPKFFKPAMPFDLMVFVTNPDGSPASKVLVVTQGS-NAKALTQDDG
RATC3   RKVLMDGVRPSSPEALVGKSLYVSVTVILHSGSDMVEAERSGIPIVTSPYQIHFT KTPKFFKPAMPFDLMVFVTNPDGSPARRVPVVTQGS-DAQALTQDDG
GUIPIG  RQVLLDGVQPSRPEALVGKSLYVSVTVILHSGSDMVEAERSGIPIVTSPYQIHFT KTPKYFKPAMPFEIMVLVTNPDGSPAPHVPVVTQGS-NVQSLTQADG
        *.**..**.  . ********** ************.****************** **** *** ***..** ********* .* * **   ..*** **

HuC3    VAKLSINTHPSQKPLSITVRTKKQELSEAEQATRTMQALPYSTVGNSNNYLHLSVLRTELRPGETLNVNFLLR MDRAHEAKIRYYTYLIMNKGRLLKAGRQV
MuC3    VAKLSINTPNSRQPLTITVRTKKDTLPESRQATKTMEAHPYSTMHNSNNYLHLSVSRMELKPGDNLNVNFHLR TDPGHEAKIRYYTYLVMNKGKLLKAGRQV
RATC3   VAKLSVNTPNNRQPLTITVSTKKEGIPDARQATRTMQAQPYSTMHNSNNYLHLSVSRVELKPGDNLNVNFHLR TDAGQEAKIRYYTYLVMNKGKLLKAGRQV
GUIPIG  VARLSINTPNTRQPLSVTVQTKKGGIPDARQAINTMQALPYTTMYNSNNYLHLSMPRTELKPGETINVNFHLR SDPNQEAKIRYYTYLIMNKGKLLKVGRQP
        **.**.**  ..**..** ***  ...  ** **.* **.*. *********. * **.**...**** ** * .**********.****.*** ***

HuC3    REPGQDLVVLPLSITTDFIPSFRLVAYYTLIGASGQREVVADSVWVDVKDSCVGSLVVKS-----GQSEDRQPVPGQQMTLKIEGDHGARV VLVAVDKGVFVL
MuC3    REPGQDLVVLSLPITPEFIPSFRLVAYYTLIGASGQREVVADSVWVDVKDSCIGTLVVKG----DPRDN--HLAPGQQTTLRIEGNQGARV GLVAVDKGVFVL
RATC3   REPGQDLVVLSLPITPEFIPSFRLVAYYTLIGANGQREVVADSVWVDVKDSCVGTLVVKG----DPRDNR-QPAPGHQTTLRIEGNQGARV GLVAVDKGVFVL
GUIPIG  REPGQALVVLPMPITKELIPSFRLVAYYTLIGASAQREVVADSVWADVRDSCVGTLVVKGGSGKDGQDKRQQHLPRQQMTLRIEGNQGARV GLVAVDKGVFVL
        ***** **** . ** . *************** ********** **.***.*.****          .    . * .* **.*** .**** ***********
```

FIG. 4B

```
HuC3    NKKNKLTQSKIWDVVEKADIGCTPGSGKDYAGVFSDAGLTFTSS-SGQQTAQRAELQCPQPAARRRRSVQLTEKRMDKVGKYP-KELRKCCEDGMRENPMRFSCQRRT
MuC3    NKKNKLTQSKIWDVVEKADIGCTPGSGKNYAGVFMDAGLAFKTS-QGLQTEQRADLECTKPAARRRRSVQLMERRMDKAGQYTDKGLRKCCEDGMRDIPMRYSCQRRA
RATC3   NKKNKLTQSKIWDVVEKADIGCTPGSGKNYAGVFMDAGLTFKTN-QGLQTDQREDPECAKPAARRRRSVQLMERRMDKAGQYTDKGLRKCCEDGMRDIPMPYSCQRRA
GUIPIG  NKKHKLTQSKIWDVVEKADIGCTPGSGKDYAGVFTDAGLSFKSSKAGLQTAQREGLDCPKPAARRRRSVQLMERRMDKAGKYKSKELRRCCEDGMRENPMQFSCQRRA

HuC3    RFISLGEACKKVFLDCCNYITELRRQHARASHLGLARSNLDEDIIAEENIVSRSEFPESWLWNVEDLKEPPKNGISTKLMNIFLKDSITTWEILAVSMSDKKGICVAD
MuC3    RLITQGENCIKAFIDCCNHITKLREQHRRDHVLGLARSELEEDIIPEEDIISRSHFPQSWLWTIEELKEPEKNGISTKVMNIFLKDSITTWEILAVSLSDKKGICVAD
RATC3   RLITQGESCLKAFMDCCNYITKLREQHRRDHVLGLARSDVDEDIIPEEDIISRSHFPESWLWTIEELKEPEKNGISTKVMNIFLKDSITTWEILAVSLSDKKGICVAD
GUIPIG  RYVSLGEACVKAFLDCCTYMAQLRQQHRREQNLGLARSDMDEDIIPEEDIISRSQFPESWLWTIEELKEPERNGISTKTMNIFLKDSITTWEILAVSLSDKKGICVAD

HuC3    PFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYRQNQELKVRVELLHNPAFCSLATTKRRHQQTVTIPPKSSLSVPYVIVPLKTGLQEVEVKAAVYHHFISDGVRK
MuC3    PYEIRVMQDFFIDLRLPYSVVRNEQVEIRAVLFNYREQQELKVRVELLHNPAFCSMATAKNRYFQTIKIPPKSSVAVPYVIVPLKIGQQEVEVKAAVFNHFISDGVKK
RATC3   PYEITVMQDFFIDLRLPYSVVRNEQVEIRAVLFNYREQEKLKVRVELLHNPAFCSMATAKKRYYQTIEIPPKSSVAVPYVIVPLKIGLQEVEVKAAVFNHFISDGVKK
GUIPIG  PFEVTVMQDFFIDLRLPYSVVRNEQVEIRAVLYNYREAQSLKVRVELLHNPAFCSLATAKKRHTQTVTIGPKSSVAVPYVLVPLKIGLQEVEVKAAVYNYFISDGVKK

HuC3    SLKVVPEGIRMNKTVAVRTLDPERLGREGVQKEDIPPADLSDQVPDTESETRILLQGTPVAQMTEDAVDAERLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDETEQWE
MuC3    TLKVVPEGMRINKTVAIHTLDPEKLGQGGVQKVDVPAADLSDQVPDTDSETRIILQGSPVVQMAEDAVDGERLKHLIVTPAGCGEQNMIGMTPTVIAVHYLDQTEQWE
RATC3   ILKVVPEGMRVNKTVAVRTLDPEHLNQGGVQREDVNAADLSDQVPDTDSETRILLQGTPVAQMAEDAVDGERLKHLIVTPSGCGEQNMIGMTPTVIAVHYLDQTEQWE
GUIPIG  TLKVVPEGMRVNKTVAIRTLNPEQLGQGGVQREEIPAADLSDQVPDTDSETKILLQGTPVAQMAEDAVDAERLKHLIITPSGCGEQNMIGMTPTVIAVHYLDQTEQWE

HuC3    KFGLEKRQGALELIKKGYTQQLAFRQPSSAFAAFVK RAPSTWLTAYVVKVFSLAVNLIAIDSQVLCGAVKWLILEKQKPDGVFQEDAPVIHQEMIGGLRNNNEKDMAL
MuC3    KFGIEKRQEALELIKKGYTQQLAFKQPSSAYAAFNN RPPSTWLTAYVVKVFSLAANLIAIDSHVLCGAVKWLILEKQKPDGVFQEDGPVIHQEMIGGFRNAKEADVSL
RATC3   KFGLEKRQEALELIKKGYTQQLAFKQPISAYAAFNN RPPSTWLTAMWSRSFSLAANLIAIDSHVLCGAVKWLILEKQKPDGVFQEDGPVIHQEMIGGFRNAKEADVSL
GUIPIG  KFGLEKRQEALNLINRGYTQQLAFKQPNWAYAAFKN RASSTWLTAYVVKVFSLAANLIGIDSEVLCGAVKWLILEKQKPDGVFQEDGPVIHQEMIGGVRTAQEADVSL

HuC3    TAFVLISLQEAKDICEEQVNSLPGSITKAGDFLEANYMNLQRSYTVAI AGYALAQMGRLKGPLLNKFLTTAKDKNRWEDPGKQLYNVEATSYALLALLQLKDFDFVPP
MuC3    TAFVLIALQEARDICEGQVNSLPGSINKAGEYIEASYMNLQRPYTVAI AGYALALMNKLEEPYLGKFLNTAKDRNRWEEPDQQLYNVEATSYALLALLLLKDFDSVPP
RATC3   TAFVLIALQEARDICEGQVNSLPGSINKAGEYLEASYLNLQRPYTVAI AGYALALMNKLEEPYLTKFLNTAKDRNRWEEPGQQLYNVEATSYALLALLLLKDFDSVPP
GUIPIG  TAFVLIALQEAKDICRAQVNNLEANINKAGDYIESRYADVRRPYTLAI AGYALALLERLNGATLQKFLNAATEKNRWEEARQKLYSVEATSYALLALLLLKDFDAVPP
```

FIG. 4C

```
HuC3    VVRWLNEQRYYGGGYGSTQATFMVFQALAQYQKDAPDHQELNLDVSLQLPSRSSKITHRI HWESASLLRSEETKENEGFTVTAEGKGQGTLSVVTMYHAKAKDQLTCNKFD
MuC3    VVRWLNEQRYYGGGYGSTQATFMVFQALAQYQTDVPDHKDLNMDVSFHLPSRSSATTFRL LWENGNLLRSEETKQNEAFSLTAKGKGRGTLSVVAVYHAKLKSKVTCKKFD
RATC3   VVRWLNDERYYGGGYGSTQATFMVFQALAQYRADVPDHKDLNMDVSLHLPSRSSPTVFRL LWESGSLLRSEETKQNEGFSLTAKGKGQGTLSVVTVYHAKVKGKTTCKKFD
GUIPIG  VVRWLNEQRYYGRGYGSTQATFMVFQALAQYQTDVPDHKDLNMEVALQLPSRSSPSKFRL VWEAGSLLRSEATKQNEGFKLTAKGKGQGTLSVVAVYYAKTKRKVVCKNFD
        ******..**** ******************. * ***..**..*. .******    *. **    ***** **.** * .** ***.******..* ** *  .   *  **

HuC3    LKVTIKPAPETEKRPQDAKNTMILEICTRYRGDQDATMSILDISMMTGFAPDTDDLKQLANGVDRYISK YELDKAFSDRNTLIIYLDKVSHSEDDCLAFKVHQYFNVELIQ
MuC3    LRVSIRPAPETAKKPEEAKNTMFLEICTKYLGDVDATMSILDISMMTGFAPDTKDLELLASGVDRYISK YEMNKAFSNKNTLIIYLEKISHTEEDCLTFKVHQYFNVGLIQ
RATC3   LRVTIKPAPETAKKPQDAKSSMILDICTRYLGDVDATMSILDISMMTGFIPDTNDLELLSSGVDRYISK YEMDKAFSNKNTLIIYLEKISHSEEDCLSFKVHQFFNVGLIQ
GUIPIG  LRVTLKPAPDTVKKPQEAKSTMILGICTRYLGDQDATMSILDISMMTGFIPDTDDLKLLATGVDRYISK YEMNKDFS-KNTLIIYLDKVSHSEEECLSFKIHQFFNVGLIQ
        *.*...***.* *.*..** .* * ***.* ** *************** *** **   *. *********. * **  .*******.*.**.*..**.**.**.*** ***

HuC3    PGAVKVYAYYNLEESCTRFYHPEKEDGKLNKLCRDELCRCAEENCFIQKSDDKVTLEERLDKACEPGVDYVYKTRLVK VQLSNDFDEYIMAIEQTIKSGSDEVQVGQQRTF
MuC3    PGSVKVYSYYNLEESCTRFYHPEKDDGMLSKLCHSEMCRCAEENCFMQQSQEKINLNVRLDKACEPGVDYVYKTELTN IKLLDDFDEYTMTIQQVIKSGSDEVQAGQQRKF
RATC3   PGSVKVYSYYNLEESCTRFYHPEKDDGMLSKLCHNEMCRCAEENCFMHQSQDQVSLNERLDKACEPGVDYVYKTKLTT IELSDDFDEYIMTIEQVIKSGSDEVQAGQERRF
GUIPIG  PGSVKVYSYYNLDETCTQFYHPEKEDGMLNKLCHKDLCRCAEENCFIQLP-EKITLDERLEKACEPGVDYVYKTKLLK MELSDDFDEYIMTIEQVIKSGSDEVQAGKERRF
        **.****.****.*.**.******.** * ***. ..**********..   ... *  **.************* *   . *  ***** *.*.* ********** *..* *

HuC3    ISPIKCREALKLEEKKHYLMWGLSSDFWGEKPNLSYIIGKDTWVEHWPEEDECQDEENQKQCQDLGAFTESMVVFGCPN
MuC3    ISHIKCRNALKLQKGKKYLMWGLSSDLWGEKPNTSYIIGKDTWVEHWPEAEECQDQKYQKQCEELGAFTESMVVYGCPN
RATC3   ISHVKCRNALKLQKGKQYLMWGLSSDLWGEKPNTSYIIGKDTWVEHWPEAEERQDQKNQKQCEDLGAFTETMVVFGCPN
GUIPIG  ISHIKCRDALHLKEGKHYLMWGLSSDLWGERPNMSYIIGKDTWVEAWPEAEECQDEENQQQCQDLGTFTENMVVFGCPN
        ** .*** **.*   *.********** ***.** *********** *** .* **.  *.**..**.*** ***.****
```

FIG. 5A

```
HIV-1      FFREDLAFLQGKAREFSSEQTRANSPTISSEQTRANSPTRRELQVWGRDNNSPSEAGADRQGTVSFNFPQITLWQRPLV
HIV-2      G----------DRGLTAPRTRRGPMQGDNRGLAAPQFSLWKRPVV
SimianIV   MPRKTSGFFRAWPMGKEAPQFPHGPDASG----ADTNCSPRGSSCGSTEELHEDGQKAEGEQRETLQGGNGGFAAPQFSLWRRPIV
ChimpIV    STKKKRLLAVWARGTPNERLHRKTGEFFRERLAFPQREARQLCAE------------QNRTNGPTDRELWVPG-GREEPGEERG-REQSISTNLPQITLWQRPLI
FelineIV   KEFGKLEGGASCSP-------------SESN--------AASSNAICTSNGGETIGFVNYNKVGTTTTLEKRPEI
RSV        MEAVIK-------------------------VISSACKTYCGKTSP----SKKEIGAMLSLLQKE--
MOLONEY    GGQGQDPPPEPRITLKVGGQP------------VTFLVDTGAQHSVLTQNPGPLSDKSAWVQGATGGKRYRWTTDRKVHLA
F-MULV     TLDDQGGQGQEPPPEPRITLKVGGQP------------VTFLVDTGAQHSVLTQNPGPLSDKSAWVQGATGGKRYRWTTDRRVHLA

HIV-1      TIKIGG----QLKEA LLDTGADDTVLEEMSLPGRWKP----KMIGGIGGFIKVRQYDQILIEIC-----GHKAIG--TVLVGPTPVN-IIGRNLLTQIGCTLNF
HIV-2      TAHIEG----QPVEV LLDTGADDSIVAGIELGSNYSP----KIVGGIGGFINTKEYKNVEIEVL-----GKRVRA--TIMTGDTPIN-IFGRNILTALGMSLNL
SimianIV   TAYIEE----QPVEV LLDTGADDSIVAGIELGPNYTP----KIVGGIGGFINTKEYKDVKIKVL-----GKVIKG--TIMTGDTPIN-IFGRNLLTAMGMSLNL
ChimpIV    PVKVEG----QLCEA LLDTGADDTVIERIQLQGLWKP----KMIGGIGGFIKVKQFDNVHIEIE-----GRKVVG--TVLVGPTPVN-IIGRNILTQLGCTLVF
FelineIV   LIFVNG----YPIKF LLDTGADITILNRRDFQVKNSIENGRQNMIGVGGGKRGTNYINVHLEIRDENYKTQCIFGNVCVLEDNSLIQPLLGRDNMIKFNIRLVM
RSV        GLLMSP-------SD LYSPGSWDPIT--AALSQR-------AMILGKSG------------ELK--------TWG---LVLGALKAA-REEQVTSEQAKFWLGL
MOLONEY    TGKVTHSFLHVPDCP YPLLGRDLLTKLKAQIHFEGSG----AQVMGPMGQPLQVLTLNIEDEHR-----LHETSKEPDVSLGSTWLS-DFPQAWAETGGMGLAV
F-MULV     TGKVTHSFLHVPDCP YPLLGRHLLTKLKAQIHFEGSG----AQVVGPMGQPLQVLTLNIEDEYR-----LHETSKGPDVPLGSTWLS-DFPQAWAETGGMGLAF

HIV-1      PISP--IETVPVKLKPGMDGPKVKQWPLTEE KIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHP----A
HIV-2      PVAK--IEPIKIMLKPGKDGPRLRQWPLTKE KIEALKEICERMEKEGQLEEAPPTNPYNTPTFAIRKKDKNKWRMLIDFRELNKVTQDFTEIQLGIPHP----A
SimianIV   PIAK--VEPIKVTLKPGKDGPKLRQWPLSKE KIIALREICERMEKDGQLEEAPPTNPYNTPTFAIKKKDKNKWRMLIDFRELNKVTQDFTEVQLGIPHP----A
ChimpIV    PISS--IETVPVKLKPGMDGPKVKQWPLSAE KIKALTEICQEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHP----A
FelineIV   AQISDKIPVVKVKMKDPNKGPQIKQWPLTNE KIEALTEIVERLEKEGKVKRADSNNPWNTPVFAIKKK-SGKWRMLIDFRELNKLTEKGAEVQLGLPHP----A
RSV        GGG-------RVSP-P---GPECIEKPATER RIDKGEEVGETTVQRD-AKMAPEET--ATPKTVG----TSCYHCGTAIG-CNCATAS--------APPP-----
MOLONEY    RQA------PLIIPLKATSTPVSIKQYPMSQE ARLGIKPHIQRLLDQG--ILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVED---IHPTVPNPYNLLS
F-MULV     RQA------PLIISLKATSTPVSIKQYPMSQE ARLGIKPHIQRLLDQG--ILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVED---IHPTVPNPYNLLS

HIV-1      GLKKK-KSVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRY QYNVLPQGWKGSPAIFQSSMTKILEPFKKQNPDIVIYQYMDDLYVGSDLEIGQHRTK
HIV-2      GLAKK-RRITVLDVGDAYFSIPLHEDFRQYTAFTLPSVNNAEPGKRY IYKVLPQGWKGSPAIFQYTMRQVLEPFRKANSDVIIIQYMDDILIASDRTDLEHDKV
SimianIV   GLAKR-RRITVLDVGDAYFSIPLDEEFRQYTAFTLPSVNNAEPGKRY IYKVLPQGWKGSPAIFQHTMRNVLEPFRKANPDVTLIQYMDDILIASDRTDLEHDRV
ChimpIV    GLKKK-KSVTVLDVGDAYFSCPLDKDFRKYTAFTIPSINNETPGVRY QYNVLPQGWKGSPSIFQSSMTKILEPFREKNPDITIYQYMDDLYVGSDLEIDQHRKK
FelineIV   GLQIK-KQVTVLDIGDAYFTIPLDPDYAPYTAFTLPRKNNAGPGRRF VWCSLPQGWILSPLIYQSTLDNIIQPFIRQNPQLDIYQYMDDIYIGSNLSKKEHKEK
RSV        PYVGSGLYPSLAGVGEQQGQGGDTPPGAE-QSRAEPGHAGQAPG---- --PALTD-WAR----VREELASTGPPVV-AMP-VVIKT------EGPAWTPLEPKL-
MOLONEY    GLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISG-QL TWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQG-
F-MULV     GLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWKDPEMGISG-QL TWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQG-
```

SUBSTITUTE SHEET (RULE 26)

FIG. 5B

```
HIV-1      IEELRQHLLRWGLTTPDKKHQK-EPPFLWMGYELHP-DKWTVQP---IVLPEKDSWTVNDIQK LVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEA
HIV-2      VLQLKELLNNLGFSTPDEKFQK-DPPYRWMGYELWP-TKWKLQK---IQLPQKEVWTVNDIQK LVGVLNWAAQIYPGIKTKHLCRLIRGKMTLTEEVQWTELA
SimianIV   VLQLKELLNSIGFSTPEEKFQK-DPPFQWMGYELWP-TKWKLQK---IELPQRETWTVNDIQK LVGVLNWAAQIYPGIKTKHLCRLIRGKMTLTEEVQWTEMA
ChimpIV    VEELRQHLLKWGFTTPDKKHQK-EPPFLWMGYELHP-DKWTVQP---IQLPEKEVWTVNDIQK LIGKLNWASQIYPGIKIKQLCKLIRGTKKLTDVVPLTPEA
FelineIV   VEELRKLLLWWGFETPEDKLQE-EPPYTWMGYELHP-LTWTIQQ---KQLDIPEQPTLNELQK LAGKINWASQAIPDLSIKALTNMMRGNQNLNSTRQWTKEA
RSV        ITRLADTVRTKGLRSPITMAEV-------EALMSSP-------L-----LPHD---VTNLMRV ILG---------PA-------PYALWMD------AWG---
MOLONEY    TRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLRE FLGTAGFCRLWIPG-FAEMAAPLYPLTKTGT-LFNWGPDQ
F-MULV     TRALLQTLGDLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLRE FLGTAGLCRLWIPG-FAEMAAPLYPLTKTGT-LFKWGPDQ

HIV-1      ELELAEN-REILKEPVHG---------VYYDPSKDLIAEIQKQGQGQWTYQIYQEP-FKNLKTGKYARMRGAHTNDVKQL TEAVQKITTESIVIWGKTPKFKL
HIV-2      EAELEEN-RIILSQEQEG---------HYYQEEKELEATVQKDQDNQWTYKIHQE--EKILKVGKYAKIKHTHTNGVKLL AQVVQKIGKEALVIG-RIPKFHL
SimianIV   EAEYEEN-KIILSQEQEG---------CYYQEGKPLEATVIKSQDNQWSYKIHQE--DKILKVGKFAKIKNTHTNGVRLL AHVVQKIGKEAIVIWGQVPRFHL
ChimpIV    ELELAEN-REIVSTPVHG---------VYYDPDKELIAEIQKQGNCQWTYQIFQEP-HKNLKTGKYARQRSAHTNDIRQL AEAVQKIATESIVIWGKTPKFRL
FelineIV   RLEVQKAKKAIEEQVQLG---------YYD-PSKELYAKLSLVGPHQISYQVYQKDPEKILWYGKMSRQKKKAENTCDIA LRACYKIREESIIRIGKEPRYEI
RSV        -VQLQTVIAAATRDPRHP--------------------AN--GQGRGERTNLNR--LKGLADGMVGNPQG--QAALLRP GELV-AITASALQAFREVARLAE
MOLONEY    QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS-KKLDPVAAGWPPCLRMVAAIAVL TKDAGKLTMGQPLVILAPHAVEA
F-MULV     QKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS-KKLDPVAAGWPPCLRMVAAIAVL TKDVGKLTMGQPLVILAPHAVEA

HIV-1      PIQKET--WET------WWTEYWQATWIPEWEFVNTPPLVKLWY--------------------QLEKEPIVGAE-TFYVDGAANRE--TKLGKAG YVTNKG
HIV-2      PVEREV--WEQ------WWDNYWQVTWIPDWDFVSTPPLVRLAF--------------------NLVGDPIPGTE-TFYTDGSCNRQ--SKEGKAG YVTDRG
SimianIV   PVEREI--WEQ------WWTDYWQVTWIPEWDFVSTPPLVRLVF--------------------NLVKEPIQGAE-TFYVDGSCNRQ--SREGKAG YVTDRG
ChimpIV    PVQKES--WEA------WWAEYWQATWIPEWEFINTPPLVKLWY--------------------SLETEPIPTTD-TYYVDGAANRE--TKTGKAG YVTDKG
FelineIV   PTSREA--WESN-----LINSPYLKAPPPEVEYIHAALNIKRALS-------------------MIKDAPIPGAE-TWYIDGGRKLG---KAAKAA YWTDTG
RSV        PAGP----WAD-------------IMQGPSESFV------------------------------DFANRLIKAVE--G-SD----LP---PSARAP VIIDCF
MOLONEY    LVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAA VTTETE
F-MULV     LVKQPPDRWLSNARMTHYQALLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRRAGAA VTTETE

HIV-1      RQKVVPLTN-TTNQKTELQAIYLALQDS-GLEVNIVTDSQYAL-------------GIIQAQPDKSESE-LVNQIIEQLIKKEKVYLAWVPAH-KG----IGG
HIV-2      RDKVKILEQ-TTNQQAELEAFAMALTDS-GPKANIIVDSQYVM-------------GIVAGQPTESENR-IVNQIIEEMIKKEAIYVAWVPAH-KG----IGG
SimianIV   RDKAKLLEQ-TTNQQAELEAFYLALADS-GPKANIIVDSQYVM-------------GIVAGQPTESESR-LVNQIIEEMIKKEAIYVAWVPAH-KG----IGG
ChimpIV    KQKIISLEN-TTNQQAELKALLLALQDS-DQQVNIVTDSQYVL-------------GIIQSQPDHSESE-LVNQIIEELIKKEKIYLSWVPAH-KG----IGG
FelineIV   KWRVMDLEG--SNQKAEIQALLLALKAG-SEEMNIITDSQYVIN-------------IILQQPDMMEG--IWQEVLEELEKKTAIFIDWVPGH-KG----IPG
RSV        RQ--KSQPD---IQ-QLIRTAPSTLTTP-GEIIKYVLDRQKTA--------------PLTDQGIAAAMSSAIQPLIMAVVNRE------RDGQ-TG----SGG
MOLONEY    VIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARG
F-MULV     VIWAKALPAGTSAQRAELIALTQALKMAAGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGNHAEARG
```

FIG. 5C

```
HIV-1      NEQVDKLV-SA GIRK---------------------------------------------------------------------------ILFLDGIDKAQDE
HIV-2      NQEVDHLV-SQ GIRQ---------------------------------------------------------------------------VLFLEKIEPAQEE
SimianIV   NQEVDHLV-SQ GIRQ---------------------------------------------------------------------------VLFLEKIEPAQEE
ChimpIV    NEQVDKLV-SA GIRK---------------------------------------------------------------------------VLFLDGIDRAQEE
FelineIV   NEEVDKLCQTM MIIEGDGILDKRSEDAGYDLLAAKEIHLLPGEVKVIPTGVKLMLPKGYWGLIIGKSSIGSKGLDVLGGVIDEGYRGEIGVIMINVSRKSITL
RSV        R--------AR GLCY---------------------------------------------------------------------------TCGSPGHYQAQCP
MOLONEY    NRMADQAARKA AITETP----------------------------------------------------------------------DTSTLLIENSSPYTSE
F-MULV     NRMADQAAREV ATRETP----------------------------------------------------------------------ETSTLLIENSAPYTRE

HIV-1      HEKY------------------------
HIV-2      HEKY------------------------
SimianIV   HEKY------------------------
ChimpIV    HERY------------------------
FelineIV   MERQKIAQLIILPCKHEVLEQGKVVMDS
RSV        KKRK------------------------
MOLONEY    HFHYTVTDIKDLTKLGAIYDKTKKYWVY
F-MULV     HFHYTVTDIKDLTKLGATYDDAKKCWVY
```

FIG. 6A

```
                                                                                       PR
HIV-1    FFREDLAFLQGKAREFSSEQTRANSPTISSEQTRANSPTRRELQVWGRDNNSPSEAGADRQGT-VSFNFPQITLWQRPLVTIKIGGQLKEALLDTGADDTVL
HIV-2                                   GDRGLTAPRTRRGPMQG-----------DNRG----LAAPQFSLWKRPVVTAHIEGQPVEVLLDTGADDSIV
MOLONEY        GGQGQDPPPEPRITLKVGGQPVTFLVDTGAQHSVLTQNPGPLSDKSAWVQGATGGKRYRWTTDRKVHLATGKVTHSFLHVPDCPYPLLGRDLLTK
F-MULV     TLDDQGGQGQEPPPEPRITLKVGGQPVTFLVDTGAQHSVLTQNPGPLSDKSAWVQGATGGKRYRWTTDRRVHLATGKVTHSFLHVPDCPYPLLGRHLLTK

                                                                                 PR
HIV-1    EEMSLPGRWKPRMIGGIG GFIKVRQYDQILIEICGHKAIG--TVLVGPTPVNIIGRNLLTQIGCTLNFPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKAL
HIV-2    AGIELGSNYSPKIVGGIG GFINTKEYKNVEIEVLGKRVRA--TIMTGDTPINIFGRNILTALGMSLNLPVAKIEPIKIMLKPGKDGPRLRQWPLTKEKIEAL
MOLONEY  LKAQIHFEGSGAQVMGPM GQPLQVLTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLI---IPLKATSTPVSIKQYPMSQEARLGI
F-MULV   LKAQIHFEGSGAQVVGPM GQPLQVLTLNIEDEYRLHETSKGPDVPLGSTWLSDFPQAWAETGGMGLAFRQAPLI---ISLKATSTPVSIKQYPMSQEARLGI

HIV-1    VEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTK WRKLVDFRELNKRTQDFWEVQLGIPHP----AGLKKK-KSVTVLDVGDAYFSVPLDEDFRKYTAFT
HIV-2    KEICEKMEKEGQLEEAPPTNPYNTPTFAIRKKDKNK WRMLIDFRELNKVTQDFTEIQLGIPHP----AGLAKK-RRITVLDVGDAYFSIPLHEDFRQYTAFT
MOLONEY  KPHIQRLLDQG--ILVPCQSPWNTPLLPVKKPGTND YRPVQDLREVNKRVED---IHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFE
F-MULV   KPHIQRLLDQG--ILVPCQSPWNTPLLPVKKPGTND YRPVQDLREVNKRVED---IHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFE

HIV-1    IPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFKKQNPDIVIYQYMD DLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQK-EPPFLWMGY
HIV-2    LPSVNNAEPGKRYIYKVLPQGWKGSPAIFQYTMRQVLEPFRKANSDVIIIQYMD DILIASDRTDLEHDKVVLQLKELLNNLGFSTPDEKFQK-DPPYRWMGY
MOLONEY  WRDPEMGISG-QLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVD DLLLAATSELDCQQG-TRALLQTLGNLGYRASAKKAQICQKQVKYLGY
F-MULV   WKDPEMGISG-QLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVD DLLLAATSELDCQQG-TRALLQTLGDLGYRASAKKAQICQKQVKYLGY

HIV-1    ELHP-DKWTVQP---IVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAE LELAENREILKEPVHG---------VYYD
HIV-2    ELWP-TKWKLQK---IQLPQKEVWTVNDIQKLVGVLNWAAQIYPGIKTKHLCRLIRGKMTLTEEVQWTELAE AELEENRIILSQEQEG---------HYYQ
MOLONEY  LLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAP-LYPLTKTGTLFNWGPDQQK AYQEIKQALLTAPALGLPDLTKPFELFVD
F-MULV   LLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGLCRLWIPGFAEMAAP-LYPLTKTGTLFKWGPDQQK AYQEIKQALLTAPALGLPDLTKPFELFVD

                 pol 3                                      pol 4                          pol 2
         PSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQK
         EEKELEATVQKDQDNQWTYKIHQE-EKILKVGKYAKIKHTHTNGVKLLAQVVQKIGKEALVIG-RIPKFHLPVER
         EKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ
         EKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDVGKLTMGQPLVILAPHAVEALVKQ
```

FIG. 6B

```
        pol 1
HIV-1   --EIWET------WWT EYWQATWIPEWEFVNTPPLVKLWYQLE---------------------KEPIVGA-ETFYVDGAANRETKLGKAG--YVTNKGRQKV
HIV-2   --EVWEQ------WWD NYWQVTWIPDWDFVSTPPLVRLAFNLV---------------------GDPIPGT-ETFYTDGSCNRQSKEGKAG--YVTDRGRDKV
MOLONEY PPDRWLSNARMTHYQA LLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWA
F-MULV  PPDRWLSNARMTHYQA LLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRRAGAAVTTETEVIWA

HIV-1   VPLTN-TTNQKTELQAIYLALQDS-GLEVNIVT DSQYAL--GIIQAQPDK------SES-ELVN-----QIIEQLIKKEKVYLAWVPAH-KG-----IGGNEQV
HIV-2   KILEQ-TTNQQAELEAFAMALTDS-GPKANIIV DSQYVM--GIVAGQPTE------SEN-RIVN-----QIIEEMIKKEAIYVAWVPAH-KG----IGGNQEV
MOLONEY KALPAGTSAQRAELIALTQALKMAEGKKLNVYT DSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMA
F-MULV  KALPAGTSAQRAELIALTQALKMAAGKKLNVYT DSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGNHAEARGNRMA

HIV-1   D-------------------------------KLVSAGIRKILFLDGIDKA
HIV-2   D-------------------------------HLVSQGIRQVLFLEKIEPA
MOLONEY DQAARKAAITETPDTSTLLIENSSPYTSEHFHYTVTDIKDLTKLGAIYDK
F-MULV  DQAAREVATRETPETSTLLIENSAPYTREHFHYTVTDIKDLTKLGATYDD
```

FIG. 7A

```
HSV-1                                                                MDSAAPALSPALTALTDQSATADLAIQIPKCP-DPER-YFYTSQCPDINHLRSLS
HSV-2                                                                 MDPAVSPASTDPLDTHASGAGAAPIPVCP-TPER-YFYTSQCPDINHLRSLS
EBV                                                                                                       MSKLLYVRDHEGFACLTVET
Human      MLSLRVPLAPITDPQ-QLQLSPLKG-----LSLVDKENTPPALSGTRVLASKTARRIFQEPTEPKTKAAA-PGVEDEPLLRENPRRFVIFPIEYHDIWQMYKKA
Vaccinia                                                                                         MEPILAPNPNRFVIFPIQYYDIWNMYKKA
Mus        MLSVRTPLATIADQQ-QLQLSPLKR-----LTLADKENTPPTLSSTRVLASKAARRIFQDSAELESKAPTNPSVEDEPLLRENPRRFVVFPIEYHDIWQMYKKA
Yeast       MPKETPSKAAADALSDLEIKDSKSNLNKELETLREENRVKSDMLKEKLSKDAENHKAYLKSHQVHRHKLKEMEKEEPLLNEDKERTVLFPIKYHEIWQAYKRA
Coli                                                                           AYTTFSQTKNDQLKEPMFFGQPVNVARYDQQKYDIFEKLIEK
H. infl.                                                                       AYTTFSQNKNDQLKEPMFFGQNVNVARYDQQKYETFEKLIEK

HSV-1      ILN-RWLETELVFVGD EEDVSKLSEG-ELSFYRFLFAFLSAADDLVTENLGG-LSGLFEQKDILHYYVEQECIEVVHSRVYNIIQLVLFHNNDQARREYVAGT
HSV-2      ILN-RWLETELVFVGD EEDVSKLSEG-ELGFYRFLFAFLSAADDLVTENLGG-LSGLFEQKDILHYYVEQECIEVVHSRVYNIIQLVLFHNNDQARRAYVART
EBV        HRN-RWFAAHIVLTKD CGCLKLLNER-DLEFYKFLFTFLAMAEKLVNFNIDE-LVTSFESHDIDHYYTEQKAMENVHGETYANILNMLFDG-DRAAMNAYAEA
Human      EAS-FWTAEEVDLSKD IQHWESLKPE-ERYFISHVLAFFAASDGIVNENLVERFSQEVQITEARCFYGFQIAMENIHSEMYSLLIDTYIK--DPKEREFLFNA
Vaccinia   EAS-FWTVEEVDISKD INDWNKLTPD-EKYFIKHVLAFFAASDGIVNENLAERFCTEVQITEARCFYGFQMAIENIHSEMYSLLIDTYVK--DSNEKNYLFNA
Mus        EAS-FWTAEEVDLSKD IQHWEALKPD-ERHFISHVLAFFAASDGIVNENLVERFSQEVQVTEARCFYGFQIAMENIHSEMYSLLIDTYIK--DPKEREYLFNA
Yeast      EAS-FWTAEEIDLSKD IHDWNNRMNENERFFISRVLAFFAASDGIVNENLVENFSTEVQIPEAKSFYGFQIMIENIHSETYSLLIDTYIK--DPKESEFLFNA
Coli       QLSFFWRPEEVDVSRD RIDYQALPEH-EKHIFISNLKYQTLLDSIQGRSPNVALLPLISIPELETWVETWAFSETIHSRSYTHIIRNIVN--DPS---VVFDD
H. infl.   QLSFFWRPEEVDVSQD RIDYAALPEH-EKHIFISNLKYQTLLDSIQGRSPNVALLPLVSIPELETWIETWTFSETIHSRSYTHIIRNIVN--DPS---IVFDD
              *  .  *          .          . . . .               .               * .*  *  .        *

HSV-1      IN-HPAIRAKVDWLEARVREC-------ASVPE KFILM----------------------ILIEGIFFAASFAAIAYLRTNNLLRVTCQSNDLISRDEAVHT
HSV-2      IN-HPAIRVKVDWLEARVREC-------DSIPE KFILM----------------------ILIEGVFFAASFAAIAYLRTNNLLRVTCQSNDLISRDEAVHT
EBV        IMADEALQAKISWLRDKVAAA-------VTLPE KILVF----------------------LLIEGIFFISSFYSIALLRVRGLMPGICLANNYISRDELLHT
Human      IETMPCVKKKADWALRWIGDKE------ATYGE RVVAF----------------------AAVEGIFFSGSFASIFWLKKRGLMPGLTFSNELISRDEGLHC
Vaccinia   IETMPCVKKKADWAQKWIHDS-------AGYGE RLIAF----------------------AAVEGIFFSGSFASIFWLKKRGLMPGLTFSNELISRDEGLHC
Mus        IETMPCVKKKADWALRWIGDKE------ATYGE RVVAF----------------------AAVEGIFFSGSFASIFWLKKRGLMPGLTFSNELISRDEGLHC
Yeast      IHTIPEIGEKAEWALRWIQDAD------ALFGE RLVAF----------------------ASIEGVFFSGSFASIFWLKKRGMMPGLTFSNELICRDEGLHT
Coli       IVTNEQIQKRAEGISSYYDELIEMTSYWHLLGE GTHTVNGKTVTVSLRELKKKLYLCLMSVNALEAIRFYVSFACSFAFAERELMEGNAKIIRLIARDEALHL
H. infl.   IVTNEEIIKRAQDISSYYDDLIRDSQLYGLYGE GTYTVDGKECVVTLRSLKKQLYLCLMSVNALEAIRFYVSFACSFAFAERRLMEGNAKIIKFIARDEALHL
           *   .  .                      *                            .*  .  *  **  .      ..   .     *.***  .*
```

FIG. 7B

```
HSV-1      TASCYIYNNYLGGHAKP--------PPDRVYGLFRQAVEIEIGFIRSQAP TDSHILSPAALAAIENYVRFSADRLLGLIHM-KPLFSAPPPDASFPLSLMSTD
HSV-2      TASCYIYNNYLGGHAKP--------EAARVYRLFREAVDIEIGFIRSQAP TDSSILSPGALAAIENYVRFSADRLLGLIHM-QPLYSAPAPDASFPLSLMSTD
EBV        RAASLLYNSMTAKADRP--------RATWIQELFRTAVEVETAFIEARGE -G---VTLVDVRAIKQFLEATADRILGDIGQ-APLYGTPPP-KDCPLTYMTSI
Human      DFACLMFKHLV--HK-P--------SEERVREIIINAVRIEQEFLTEALP VK---LIGMNCTLMKQYIEFVADRLMLELGF-SKVFRVENP-FDF-MENISLE
Vaccinia   DFACLMFKHLL--HP-P--------SEETVRSIITDAVSIEQEFLTAALP VK---LIGMNCEMMKTYIEFVADRLISELGF-KKIYNVTNP-FDF-MENISLE
Mus        DFACLMFKHLV--HK-P--------AEQRVREIITNAVRIEQEFLTEALP VK---LIGMNCTLMKQYIEFVADRLMLELGF-NKIFRVENP-FDF-MENISLE
Yeast      DFACLLFAHLK--NK-P--------DPAIVEKIVTEAVEIEQRYFLDALP VA---LLGMNADLMNQYVEFVADRLLVAFGN-KKYYKVENP-FDF-MENISLA
Coli       TGTQHMLNLLRSGADDPEMAEIAEECKQECYDLFVQAAQQEKDWADYLFR DGS--MIGLNKDILCQYVEYITNIRMQAVGLDLPFQTRSNPIPWINTWLVSDN
H. infl.   TGTQHILNIMAAGQDDPEMAEIAEECKQEAYDLFVAAAEQEKAWADYLFK DGS--MIGLNRDILVQYVEYITNIRMQAVGLPLPFQTRSNPIPWINAWLVSDN
             . .      *           .   *   *                    .     . ..   .    .          *            ..

HSV-1      KHTNFFECRSTSYAGAVVNDL
HSV-2      KHTNFFECRSTSYAGAVVNDL
EBV        KQTNFFEQESSDYTMLVVDDL    ;;
Human      GKTNFFEKRVGEYQRMGVMSSP---TENSFTLDADF
Vaccinia   GKTNFFEKRVGEYQKMGVMSQ----EDNHFSLDVDF
Mus        GKTNFFEKRVGEYQRMGVMSNS---TENSFTLDADF
Yeast      GKTNFFEKRVSDYQKAGVMSKSTKQEAGAFTFNEDF
Coli       VQVAPQEVEVSSYLVGQIDSEVDTDDLSNFQL
H. infl.   VQVAPQEVEVSSYLVGQIDSKVDTNDFDDFSL
           .   *      *      .
```

FIG. 8A

```
HSV TYPE1/F                                                  MDADGASPPPPRPAGGP------------K------NTPAAPPLYATGRLSQAQLMPSPPMPVPPA
HSV TYPE2/HG52                                      MDLLVDDLFADRDGVSPPPPRPAGGP------------K------NTPAAPPLYATGRLSQAQLMPSPPMPVPPA
BvHV Type1/P8-2                                            MSGRIKTAGRALASQCGGAAAATMDPYDAIEAFDDSLLGSPLAAGPLYDGPSPARFALP-PPRPAPLA
Var-ZosV/Dumas                                             MECNLG-TEHPSTDTWNRS-KTEQAVVDAFDESLF--GDVASDIGFETSLYSHAVKTAPSPPWVASPK
EqHV Type4                                             MAANIAMFADIEDYDDTRSCEYGYG-TCELMDVDGVVASFD--EGMLSASESIYSSPAQKRLALP-PPKATSPT
EqHV Type1/AB4P    MCLLHISLPYLSCALLPGWYFDARPAASIVMFAAAEENDDPYPGKSGYNDTCELMDMDGAVASFD--EGMLSAIESVYSIPTKKRLALP-PPKAASPG

HSV TYPE 1/F       ALFNRLLDDLGFSAGPALCTML DTWNEDLFSALPTNADLYRECKFLSTLPSDVVEWG----DAYVPERAQIDIRAHGDVAFPTLPATRDGLGLYYEALS
HSV TYPE2/HG52     ALFNRLLDDLGFSAGPALCTML DTWNEDLFSGFPTNADMYRECKFLSTLPSDVIDWG----DAHVPERSPIDIRAHGDVAFPTLPATRDELPSYYEAMA
BvHV Type1/P8-2    ALLERMQAELGFPDGPALLRAM ERWNEDLFSCLPTNADLYADAALLSADADAVVGAMY---LAVPGDAERLDLNAHANQPLPAPPASEEGLPEYVAGVQ
Var-ZosV/Dumas     ILYQQLIRDLDFSEGPRLLSCL ETWNEDLFSCFPINEDLYSDMMVLSPDPDDVIST-----VSTKDHVEMFNLTTRGSVRLPSPPKQPTGLPAYVQEVQ
EqHV Type4         ALYQRLQAELGFPEGQAMLFAM EKWNEDMFSAIPVHVDLYTEIALLSTSVNEVVKAGLDSLPIPTNYIPEVDLNAHGSEPFPEVPALEDELETYVISAQ
EqHV Type1/AB4P    ALYQRLQGELGFPEGQTLLSAM EKWNEDMFSALPGHVDLYTEIALLSTSVDEVVRAGLDSLPTPSHYSPEVDLNAHGDEPFPEVPALEDDLEIYVISAQ

HSV TYPE1/F        RFFHAELRAREESYRTVLANFCSALYRYLRASVRQLHRQAHMR GRDRDLGEMLRATIADRYYRETARLARVLFLHLYLFLTREILWAAYAEQMMRPDLF
HSV TYPE2/HG52     QFFRGELRAREESYRTVLANFCSALYRYLRASVRQLHRQAHMR GRNRDLREMLRTTIADRYYRETARLARVLFLHLYLFLSREILWAAYAEQMMRPDLF
BvHV Type1/P8-2    AHFLAELRAREERYAGLFLGYCRALLQHLRATAARGRGAAGAG ----AQADRLRQLVAARYYREASRLARLAFAHMYVATAREVSWRLHSQQSQAQGVF
Var-ZosV/Dumas     DSFTVELRAREEAYTKLLVTYCKSIIRYLQGTAKRTTIGLNIQ NPDQKAYTQLRQSILLRYYREVASLARLLYLHLYLTVTREFSWRLYASQSAHPDVF
EqHV Type 4        RFYLSELRAREEHYSRLLRGYCVALLHYLYGSAKRQLRGAGS- --DSALMHKFKQVVRDRYYRETANLARLLYLHLYISVTREVSWRLHASQVVNQGIF
EqHV Type1/AB4P    RFYLSELRTREEHYARLLRGYCVALLHYLYGSAKRQLRGSGS- --DASLMHKFKQVVRDRYYREAANLARLLYLHLYVSVTREVSWRLHASQVINQGVF

HSV TYPE 1/F       DCLCCDLESWRQLAGLFQPFMFVNGALTVRGVPIEARRLRELNHIREHLNLPLVRSAATEEPGA PLTTPPTLHGNQARASGYFMVLIRAKLDSYSSFTT
HSV TYPE2/HG52     DGLCCDLESWRQLACLFQPLMFINGSLTVRGVPVEARRLRELNHIREHLNLPLVRSAAAEEPGA PLTTPPVLQGNQARSSGYFMLLIRAKLDSYSSVAT
BvHV Type1/P8-2    VSLYYAWPQRRQFTCLFHPVLFNHGVVALEDGFLDAAELRRLNYRRRELGLPLVRAGLVEVEVG PLVEEPPFSGSLPRALGFLNYQVRAKMGAPAEAGG
Var-ZosV/Dumas     AALKFTWTERRQFTCAFHPVLCNHGIVLLEGKPLTASALREINYRRRELGLPLVRCGLVEENKS PLVQQPSFSVHLPRSVGFLTHHIKRKLDAYAVKHP
EqHV Type4         VSLHYTWPQRRKFECLFHPVLFNHGVVILENDPLEFNDLQRINYRRRELGLPLIRAGLIEEENL PLESEPTFSGKLPRTIGFLTHQIRTKMEAYSNAHP
EqHV Type1/AB4P    VSLHYFWAQRRKFECLFHPVLFNHGVVILENDPLEFHDLQRINYRRRELGLPLIRAGLIEEENS PLEAEPLFSGKLPRTIGFLTHQIRTKMEAYSDAHP
```

FIG. 8B

```
HSV TYPE 1/F       SPSEAVMREHAYSRAR-TKNNYGSTIEGLLDLPDDD-APEEAGLAAPRLSFLPAGHTRR---------LSTAPPTDVSLGDELHLD GEDVAMAHA
HSV TYPE2/HG52     SEGESVMREHAYSRGR-TRNNYGSTIEGLLDLPDDDDAPAEAGLVAPRMSFLSAGQRPRRL-------STTAPITDVSLGDELRLD GEEVDMTPA
BvHV Type1/P8-2    GWRR---SGSTRTRGRAARSTTGRLQRPCCGPRR----RAKCCRATPRQRLRARGEPRHTSGSGAFSQGRRPGRVCRLGWACKAR SGPARGGPG
Var-ZosV/Dumas     QEPRHVRADHPYAKVVENRN-YGSSIEAMILA-------PPSPSEILP--------GDPPR-------------PPTCG-----FLTR
EqHV Type4         STPLFPLAEHSYSKRIDGRLSYGTTAEAMMD--------PPSPSAVLP--------GDPVP-------------PLTVG-----IRQT AETLALPSN
EqHV Type1/AB4P    ATPLFPLAEHSYSKRIGGRLSYGTTTEAMMD--------PPSPSAVLP--------GDPVP-------------PLTVG-----VRQT AATLAIPSN
                          ..    .    *                              *         *                  *

HSV TYPE 1/F       DALDDFDLDMLGDGDSPGPGFTPHDSAPYGA---LDMADFEFEQMFTDALGIDEYGG
HSV TYPE2/HG52     DALDDFDLEMLGDVESPSPGMT-HDPVSYGA---LDVDDFEFEQMFTDAMGIDDFGG
BvHV Type1/P8-2    PSPVRSGLGLSRARGSPGPGPACGGPSRARGGRRRASPANPFGGTYDALLGDRLNQLLDF
Var-ZosV/Dumas
EqHV Type4         LTLQSMETDVLDYS---------------------SISGDELNQMFDI
EqHV Type1/AB4P    LTLQSMETDGLDYS---------------------SMTGDELNQMFDI
```

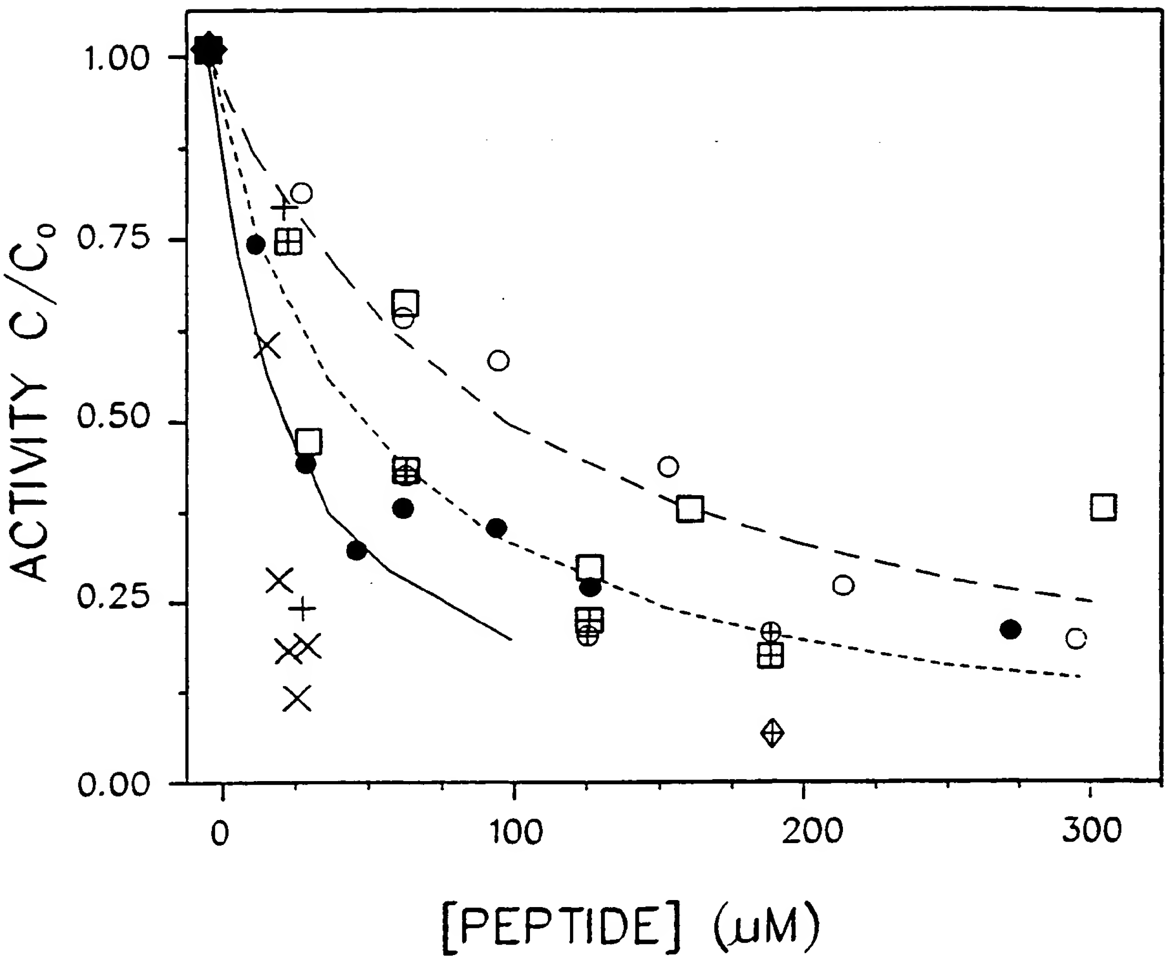

FIG. 9

FIG. 10

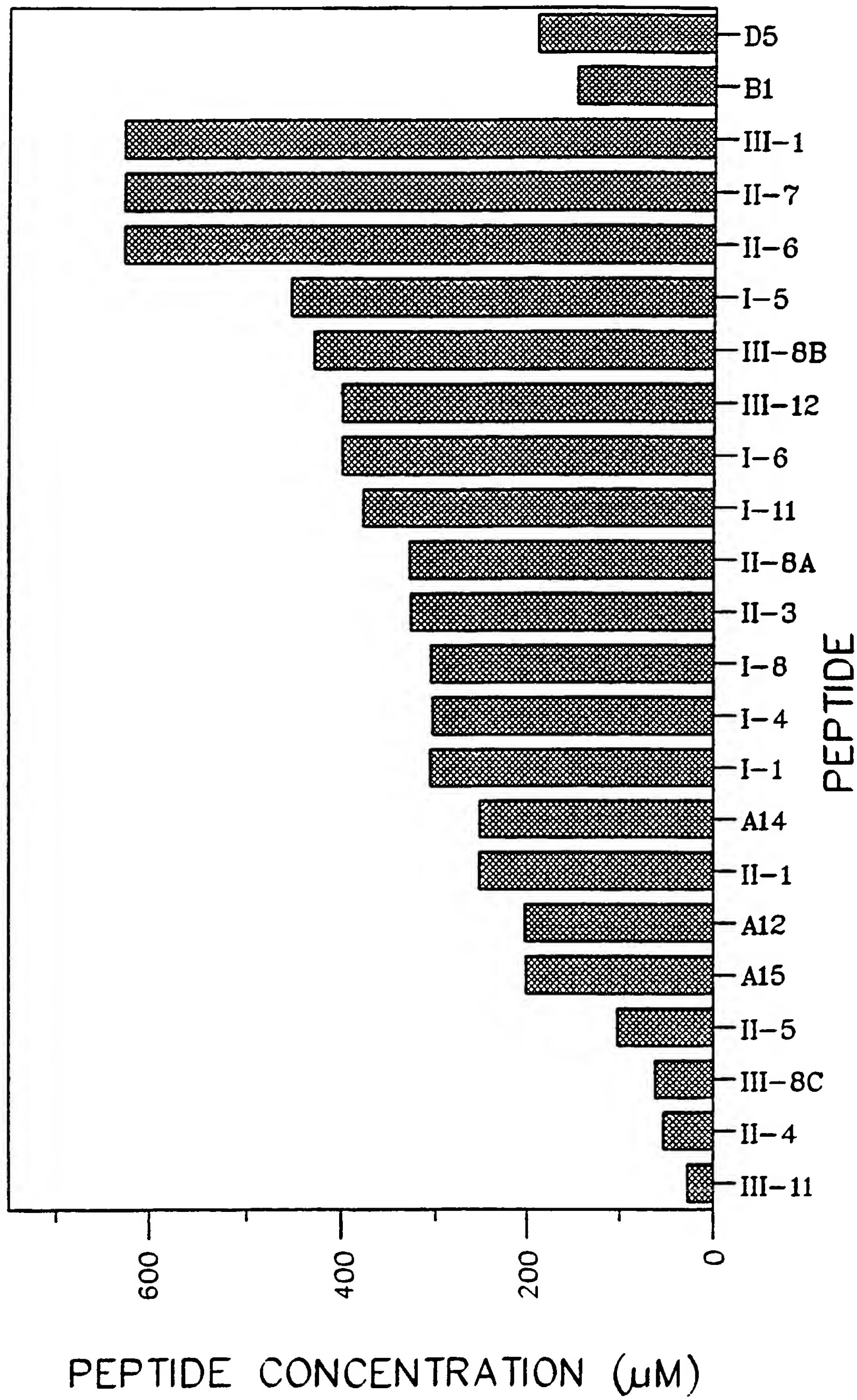
D5
B1
III-1
II-7
II-6
I-5
III-8B
III-12
I-6
I-11
II-8A
II-3
I-8
I-4
I-1
A14
II-1
A12
A15
II-5
III-8C
II-4
III-11
PEPTIDE
600
400
200
0
PEPTIDE CONCENTRATION (μM)

FIG. 11A

```
humC2   MGP--LMVLFCLLFLYPGLADS--------APSCPQNVNISGGTFTLSHGWAPGSLLTYSCPQGLYPSPAS-RLCKSSGQWQTPG-------ATRSLSKAVCKPV
musC2   MAP--LLALFYLLQLGPGLA----------ALFCNQNVNITGGNFTLSHGWAPGSLLIYSCPLGRYPSPAW-RRCQSNGQWLTPRSSSHHTLRSSRMVKAVCKPV
humBf   MGSNLSPQLCLMPFILGLLSGGVTTTPWSLARPQGSCSLEGVEIKGGSFRLLQ---EGQALEYVCPSGFYPYPVQTRTCRSTGSWSTLKTQ-----DQKTVRKAECRAI
musBf   MESPQLCLVLLVLGFSSGGVSATPVLEARPQVSCSLEGVEIKGGSFQLLQ---GGQALEYLCPSGFYPYPVQTRTCRSTGSWSDLQTR-----DQKIVQKAECRAI
zebBf   MTSMECGLRLKWLILALICPLTAGAP----SREGSCPEENLDIAGGSFTLSNGYSDGSYLQYICPDNHYPSISS-RRCQ-FGVWTPKAS---------SRKKAECKKI
        .    *    .    ... * ** * *    * * * ** . **    * *.  * *    ** *. .

humC2   RCPAPVSFENG IYTPRLGSYPVGGNVSFECEDGFILRGSPVRQCRPNGMWDGETAVCDNGAGHCPNPGISLGAVRTGFRFGHGDKVRYRCSS-NLVLTGSSERECQGN
musC2   RCLAPSSFENG IYFPRLVSYPVGSNVSFECDEDFTLRGSPVRYCRPNGLWDGETAVCDNGASHCPNPGISVGTARTGLNFDLGDKVRYRCSSSNMVLTGSAERECQSN
humBf   HCPRPHDFENG EYWPRSPYYNVSDEISFHCYDGYTLRGSANRTCQVNGRWSGQTAICDNGAGYCSNPGIPIGTRKVGSQYRLEDSVTYHCSR-GLTLRGSQRRTCQEG
musBf   RCPRPQDFENG EFWPRSPFYNLSDQISFQCYDGYVLRGSANRTCQENGRWDGQTAICDDGAGYCPNPGIPIGTRKVGSQYRLEDIVTYHCSR-GLVLRGSQKRKCQEG
zebBf   TCPNPRVLENG EVAPYQERYYINDVTTYSCSSDYKFRGSKVRVCQPNGKWNGSTPICGRDSDHCPDPGVPPGSSRTGSIFNIDDEVTYHCDS-PLTLIGSKVRSVWMY
        *  *  ***  *  *  .  .. *  . *** * *. ** * * * * .*  .  * **. *. . *  .  * * *.*  . * ** *

humC2   GVWSGTEPICRQPYSYDFPEDVA PALGTSFSH-MLGATNPTQ-KTKESLGRKIQIQRSGHLNLYLLLDCSQSVSENDFLIFKESASLMVDRIFSFEINVSVAIITFAS
musC2   GVWSGSEPICRQPYSYDFPEDVA SALDTSLTN-LLGATNPTQNLLTKSLGRKIIIQRSGHLNLYLLLDASQSVTEKDFDIFKKSAELMVERIFSFEVNVTVAIITFAS
humBf   GSW[illegible]TEPSCQDSFMYDTPQEVA EAFLSSLTETIEGVDAEDGHGPGEQQKRKIVLDPSGSMNIYLVLDGSDSIGASNFTGAKKCLVNLIEKVASYGVKPRYGLVTYAT
musBf   GSW[illegible]TEPSCQDSFMYDSPQEVA EAFLSSLTETIEGADAEDGHSPGEQQKRKIVLDPSGSMNIYLVLDGSDSIGSSNFTGAKRCLTNLIEKVASYGVRPRYGLLTYAT
zebBf   GQWSGTEPQCYADFTYDPAMEAA EAFGNSLTTTLTVQQGFED----DQHGKKISLDRGGKLDIYIAVDASDSIDPKDFDKAKKIIKTLIEKISYYEVSPNYEILMFAT
        * ***.** *  . **  . * *  * .  .    .** .  * . .*. .* * *. *  *  ..... .  .. .*.

                                                          E1                                     E2
humC2   EPKVLMSVL----NDNSRDMTEVISSLENANYKDH ENGTGTNTYAALNSVYLMMNNQMRLLGMETMAWQEIRHAIILLTDGKSNMGGSPKTAVDHIREILNINQK---
musC2   QPKTIMSIL----SERSQDVTEVITSLDSASYKDH ENATGANTYEVLIRVYSMMQTQMDRLGMETSAWKEIRHTIILLTDGKSNMGDSPKKAVTRIRELLSIEQN---
humBf   YPKIWVKVS----EADSSNADWVTKQLNEINYEDH KLKSGTNTKKALQAVYSMMSWPDDVP--P-EGWNRTRHVIILMTDGLHNMGGDPITVIDEIRDLLYIGKDRKN
musBf   VPKVLVRVS----DERSSDADWVTEKLNQISYEDH KLKSGTNTKRALQAVYSMMSWAGDAP--P-EGWNRTRHVIIIMTDGLHNMGGNPVTVIQDIRALLDIGRDPKN
zebBf   DVDQIVKMRDFKTNEKARKILKIFEDLDNFNYDKK GDRTGTNIAKLYLKILDSMSLEQVQN--K-EDFLQTQHVIIVFTDGQANMGGNPKPKVDLIKNLVIKNNAS--
        . .  .  .  *  *  .  .*.*  .  *  .* **. *** *** *  .  *. ..

humC2   -RNDYLDIYAIGVGKLDVDWRELNELGSKKDGERHAFILQDTKALHQ VFEHMLDVSKLTDTICGVGNMSANASDQERTPWHVTIK---P-KSQETCRGALISDQWVLT
musC2   -RDDYLDIYAIGVGKLDVDWKELNELGSKKDGERHAFILQDAKALQQ IFEHMLDVSKLTDTICGVGNMSANASDQERTPWQVTFK---P-KSKETCQGSLISDQWVLT
humBf   PREDYLDVYVFGVGPL-VNQVNINALASKKDNEQHVFKVKDMENLED VFYQMIDESQ-SLSLCGMVWEHRKGTDYHKQPWQAKISVIRPSKGHESCMGAVVSEYFVLT
musBf   PREDYLDVYVFGVGPL-VDSVNINALASKKDNEHHVFKVKDMEDLEN VFYQMIDETK-SLSLCGMVWEHKKGNDYHKQPWQAKISVTRPLKGHETCMGAVVSEYFVLT
zebBf   -RENKLDLYVFGVGKD-VKKEDMNGLVSEKKDERHFFKLPDLDEVQN TFDLMLDDST-VVGLCGMQQNYDGSNKRSAYPWLAQLS--IAQSQISDCMGSLVTSRYILT
        *  **.*  ***  *  .* * * * *.* * * * .  *  .  * *.* .  .**.  **  * *.... .**
```

FIG. 11B

```
                 E4
humC2  AAHCFRDG--NDHSLWRVNVGDPKSQWGKELLIEKAVISPGFDVFAKKNQGILEFYGDD IALLKLAQKVKMSTHARPICLPCTMEANLALRRPQGS-TCRDHENE
musC2  AAHCFHDIQMEDHHLWRVNVGDPTSQHGKEFLVEDVIIAPGFNVHAKRKQGISEFYADD IALLKLSRKVKMSTHARPICLPCTVGANMALRRSPGS-TCKDHETE
humBf  AAHCFTVD--DKEHSIKVSVGGEK----RDLEIEVVLFHPNYNINGKKEAGIPEFYDYD VALIKLKNKLKYGQTIRPICLPCTEGTTRALRLPPTT-TCQQQKEE
musBf  AAHCFMVD--DQKHSIKVSVGGQR----RDLEIEEVLFHPKYNINGKKAEGIPEFYDYD VALVKLKNKLKYGQTLRPICLPCTEGTTRALRLPQTA-TCKQHKEQ
zebBf  AAHCFKEG----DTPDKITVYLEKN---TDVKVEKVFIHPNYSLTAKQSIGIKEFYDFD VALLQLKTPVKMSVNLRPICLPCTKETNRALKLSDSQGTCEKHEQI
       *****          .. *         . .*      *  . . *.  ** *** *.**..*  .*   ********  .. **.      **  .

humC2  LLNKQSVPAHFVALNG-------SKLNINLKMGVEWTSCAEVVSQEKTMFPNLTDVREVVTDQFLCSGTQE--- DESPCKGESGGAVFLERRFRFFQVGLVSWG
musC2  LLSQQKVPAHFVALNG-------NRLNINLRTGPEWTRCIQAVSQNKNIFPSLTNVSEVVTDQFLCSGMEEE-- DDNPCKGESGGAVFLGRRYRFFQVGLVSWG
humBf  LLPAQDIKALFVSEEEKK----LTRKEVYIKNGDKKGSCERDAQYAP-GYDKVKDISEVVTPRFLCTGGVSPYA DPNTCRGDSGGPLIVHKRSRFIQVGVISWG
musBf  LLPVKDVKALFVSEQGKS----LTRKEVYIKNGDKKASCERDATKAQ-GYEKVKDASEVVTPRFLCTGGVDPYA DPNTCKGDSGGPLIVHKRSRFIQVGVISWG
zebBf  LLSNELVDAAFTSKMDMEKRSPRKIRRITVKLGKYLDACVEDAKKAK--ESKWQMRRRQLQKISCGSGGNQPQR DDVSCKGESGGATHVDKYGRLIQIGVVSWG
       **   . * * .            . .. *     *                  .       .*       *   *.*.***   . .  *  *.*..***

humC2  LYNPCLGSADKNSRKRAPRSKVPPPRDFHINLFRMQPWLRQHLGD-VLNFLPL
musC2  LFDPCHGSSNKNLRKKPPRG--VLPRDFHISLFRLQPWLRQHLDG-VLDFLPL
humBf  VVDVCK---NQKRQKQVPAH----ARDFHINLFQVLPWLKEKLQDEDLGFL
musBf  VVDVCR---DQRRQQLVPSY----ARDFHINLFQVLPWLKDKLKDEDLGFL
zebBf  VKNLCS--KKRNLMQFSVSD----SRDYHINPF
       .  *     .  .          **.** *
```

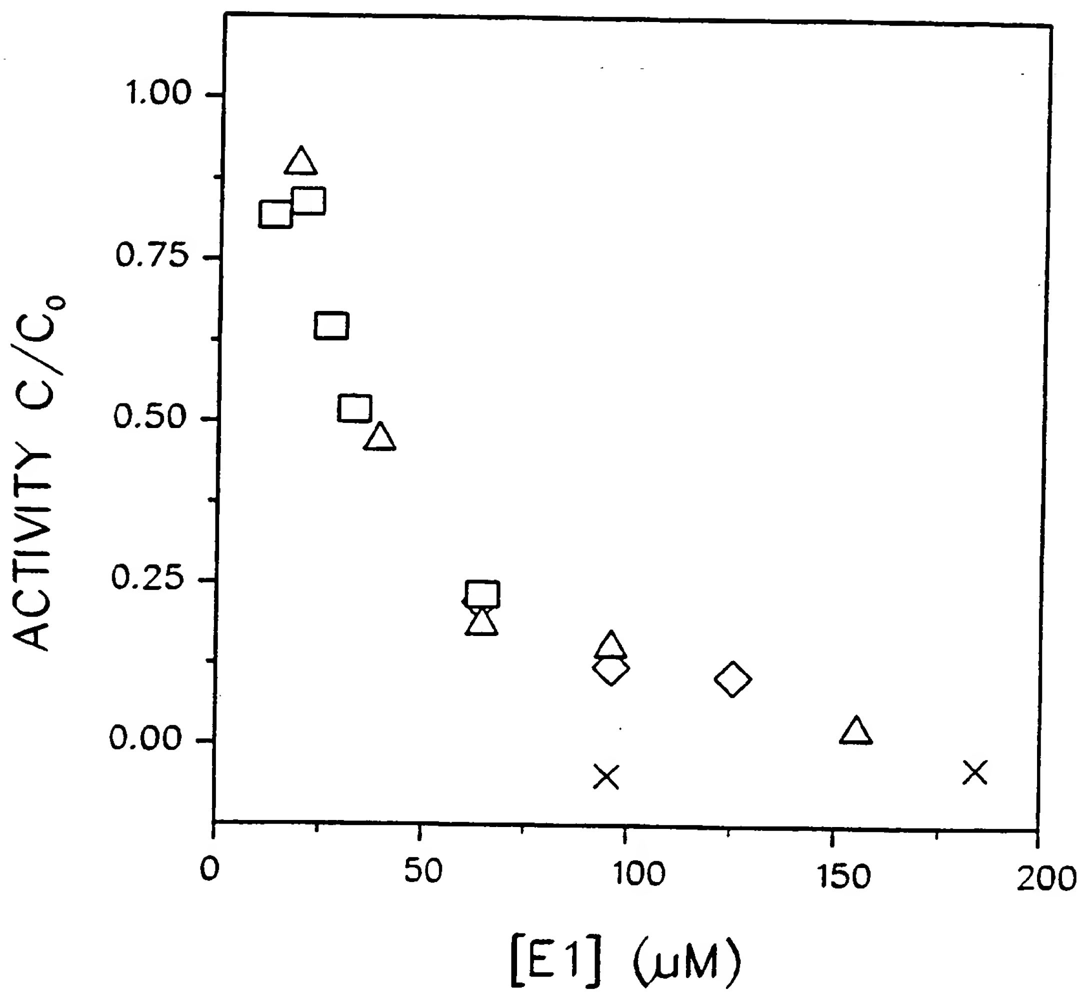

FIG. 12

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 96/10958

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K14/47 C07K16/18 A61K38/17 G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 89, no. 17, 1 September 1992, WASHINGTON US, pages 8125-8129, XP002017958 P MATHIAS ET AL.: "Mutants of complement component C3 cleaved by the C4-specific C1s-protease" see the whole document ---	18-29
X	EP,A,0 305 615 (IMMUNETECH) 8 March 1989 see the whole document ---	18-29
X	EP,A,0 312 645 (PROGEN) 26 April 1989 see the whole document ---	18-29
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

2

Date of the actual completion of the international search	Date of mailing of the international search report
8 November 1996	03.12.96

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Masturzo, P

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 86, no. 14, July 1989, WASHINGTON US, pages 5575-5579, XP002017959 R T OGATA ET AL.: "Murine complement component C4 and sex-limited protein; identification of amino acid residues essential for C4 function " see the whole document ---	18-29
X	BIOCHEMISTRY, vol. 30, no. 15, 16 April 1991, EASTON, PA US, pages 3603-3612, XP000147256 J A EMBER ET AL.: "Designing synthetic superagonists of C3a anaphylatoxin" see the whole document ---	18-29
A	BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS, vol. 194, no. 2, 30 July 1993, ORLANDO, FL US, pages 595-600, XP000382180 H J SCHRAMM ET AL.: "The inhibition of HIV-1 protease by interface peptides" cited in the application see the whole document ---	10-16
A	CHEMICAL ABSTRACTS, vol. 118, no. 23, 7 June 1993 Columbus, Ohio, US; abstract no. 228527t, B SIBANDA & J THORNTON: "Accommodating sequence changes in beta-hairpin in proteins" page 419; XP002017963 see abstract & J. MOL. BIOL., vol. 229, no. 2, 1993, pages 428-447, cited in the application ---	1-16

-/--

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CHEMICAL ABSTRACTS, vol. 116, no. 25, 22 June 1992 Columbus, Ohio, US; abstract no. 250716j, S PASCARELLA & P ARGOS: "Analysis of insertions/deletions in protein structures" page 310; XP002017964 see abstract & J. MOL. BIOL., vol. 224, no. 2, 1992, pages 461-471, cited in the application ---	1-16
X	CHEMICAL ABSTRACTS, vol. 107, no. 11, 14 September 1987 Columbus, Ohio, US; abstract no. 94949u, R BURGER ET AL.: "Functional analysis and quantification of the complement C3-derived anaphylatoxin C3a derived anaphylatoxin C3a with a monoclonal antibody" page 542; XP002017965 see abstract & CLIN. EXP. IMMUNOL. , vol. 68, no. 3, 1987, pages 703-711, ---	23
P,X	JOURNAL OF IMMUNOLOGY, vol. 155, no. 5, 1 September 1995, BALTIMORE US, pages 2642-2651, XP002017962 R T OGATA & P J LOW: "Complement component C5; engineering of a mutant that is specifically cleaved by the C4-specific C1s protease" see the whole document -----	1-30

2

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 96/10958

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
Reason: As claims 26-28 refer to non existant claims 47-49, the Search Division has read them as respectively dependent on claims 18 and 23.

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searches without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 96/10958

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-305615	08-03-89	US-A- 4692511	08-09-87
EP-A-312645	26-04-89	DE-D- 3787470	21-10-93